

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
CM nucleic - nucleic search, using sw model
Run on: March 10, 2004, 09:18:38 ; Search time 9430.35 Seconds
(without alignments)
18036.975 Million cell updates/sec

Title: US-10-084-817-25
Perfect score: 5696
Sequence: 1 gactacgtgactgcagg.....ggtttcttaaaaaaa 5696
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hct:*
9: gb_estcl:*
10: gb_estd2:*
11: gb_hct:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_nam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gssl:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	3369.4	59.2	3371 29 AY398790
2	2486.2	43.6	3371 29 AY398792
3	2408.8	42.3	3273 29 AY398791
4	1510.8	26.5	4109 11 AK080686

C	5	1416.6	24.9	3530	11	AK089315
	6	1360.8	23.9	3412	29	AY405252
	7	1219	21.4	3240	29	AY405254
	8	854	15.0	939	13	BQ933301
	9	791.6	13.9	891	13	BQ925776
	10	790.2	13.9	896	13	BX390187
C	11	760	13.3	778	14	CA414667
	12	753	13.2	763	14	CK000113
	13	744.4	13.1	937	13	BQ934054
	14	736	12.9	904	13	BQ940133
C	15	735.6	12.9	750	13	BUG18192
C	16	726.8	12.8	918	13	BX328968
	17	725.6	12.8	3401	29	AY405253
	18	725.6	12.7	746	13	BX114562
C	19	723.6	12.7	756	12	BQ008487
C	20	717	12.6	757	14	CA449275
	21	714	12.5	794	12	BI758564
C	22	705.4	12.4	962	13	BX328967
	23	703.4	12.3	876	9	AU121101
C	24	694	12.2	712	14	CA420699
	25	691.6	12.1	721	12	BM718951
C	26	689.8	12.1	707	13	BU627598
C	27	688.2	12.1	793	9	AI346330
C	28	686	12.0	701	14	CA446830
C	29	680.4	11.9	708	12	BM997135
	30	674.8	11.8	748	12	BG680063
C	31	668.2	11.7	714	14	CA428685
	32	667.8	11.7	752	14	CD511356
C	33	664.8	11.7	706	10	BF109743
C	34	662.8	11.6	718	9	AI453687
	35	650	11.4	660	9	AA149866
C	36	648.2	11.4	660	12	BM988626
C	37	641.6	11.3	693	10	BE673014
C	38	640.8	11.2	652	14	CA425104
C	39	640.4	11.2	660	14	CA431209
	40	640.2	11.2	651	12	BG939270
C	41	637.4	11.2	673	12	BQ002730
	42	636.6	11.2	672	14	CD512572
C	43	634.2	11.1	706	10	AW190828
	44	633.8	11.1	661	12	BM272400
C	45	630.8	11.1	655	12	BM971657

ALIGNMENTS

RESULT 1
AY398790
LOCUS
DEFINITION
Homo sapiens THBS2 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
ACCESSION
AY398790
VERSION
AY398790.1
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3371)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
2 (bases 1 to 3371)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
TITLE

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

Source Location/Qualifiers

1..3371

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..3371

/gene="THBS2"

/locus_tag="HCM0008"

ORIGIN

Query Match 59.2%; Score 3369.4; DB 29; Length 3371;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3370; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 132 ATGGTCTGGAGGCTGCTCCCTGCTGGCTCTGTGGGTGGCCAGCAGCAAGCTGGTCAC 191

Db 1 ATGGTCTGGAGGCTGCTCCCTGCTGGCTCTGTGGGTGGCCAGCAGCAAGCTGGTCAC 60

QY 192 CAGGCAAAAGACACGACCTTCGACCTTTTCAGTATCAGCAACATCAACCGCAAGCCATT 251

Db 61 CAGGCAAAAGACACGACCTTCGACCTTTTCAGTATCAGCAACATCAACCGCAAGCCATT 120

QY 252 GCGCCCAAGCAGTTCCGCGGCGCGACCCCGGCTGCGGCTTACCGCTTCGTGGGCTTT 311

Db 121 GCGCCCAAGCAGTTCCGCGGCGCGACCCCGGCTGCGGCTTACCGCTTCGTGGGCTTT 180

QY 312 GACTATATCCACCGGTGAACGAGATGACCTCAGCAAGATCACAAGATCATCGGGCAG 371

Db 181 GACTATATCCACCGGTGAACGAGATGACCTCAGCAAGATCACAAGATCATCGGGCAG 240

QY 372 AAGGAGGCTTCTCTCCAGGCGCCAGCTCAAGCAGGACGCAAGTCCAGGGGCAAGCTG 431

Db 241 AAGGAGGCTTCTCTCCAGGCGCCAGCTCAAGCAGGACGCAAGTCCAGGGGCAAGCTG 300

QY 432 TTGGCTCTGGAGGCGCCGCTCTCTCCAGAGGAGTTCGAGATCGTCTCCAAAGGCCCC 491

Db 301 TTGGCTCTGGAGGCGCCGCTCTCTCCAGAGGAGTTCGAGATCGTCTCCAAAGGCCCC 360

QY 492 GCGGACACGCTGATCTCACTACTGATTAACGCGACCGCGCTGCTCCCTGGAG 551

Db 361 GCGGACACGCTGATCTCACTACTGATTAACGCGACCGCGCTGCTCCCTGGAG 420

QY 552 GAGCTCGGCTGGCTGACTCGCAGTGAAGAACGTCACTGTCAGAGTGGCTGGCGAGACC 611

Db 421 GAGCTCGGCTGGCTGACTCGCAGTGAAGAACGTCACTGTCAGAGTGGCTGGCGAGACC 480

QY 612 TACAGCTTGACGCTGGGCTGGACCTCATAGACAGCTTCGCTCTGGAGGCGCTTCTAC 671

Db 481 TACAGCTTGACGCTGGGCTGGACCTCATAGACAGCTTCGCTCTGGAGGCGCTTCTAC 540

QY 672 GAGCACCCTGAGGCGGAAAAGAGCGGATGACCTGGCGCAAGGCTCTGCCAGAGAGAT 731

Db 541 GAGCACCCTGAGGCGGAAAAGAGCGGATGACCTGGCGCAAGGCTCTGCCAGAGAGAT 600

QY 732 CACTTCAGGGGTTTCTTCAGAAAGTCCACCTAGTGTGTTTGAACCTCTGTGGAAGATTT 791

Db 601 CACTTCAGGGGTTTCTTCAGAAAGTCCACCTAGTGTGTTTGAACCTCTGTGGAAGATTT 660

QY 792 CTAAGCAAGAGGGTTGCGCAGAGGCGGAGCTGAGATCAACGCCATCAGTGAGAAC 851

Db 661 CTAAGCAAGAGGGTTGCGCAGAGGCGGAGCTGAGATCAACGCCATCAGTGAGAAC 720

QY 852 ACAGAGACGCTGGGCTGGGCTCGCATGTCAACCGAGTACGTGGGCGCCAGCTCAGAG 911

Db 721 ACAGAGACGCTGGGCTGGGCTCGCATGTCAACCGAGTACGTGGGCGCCAGCTCGGAG 780

QY 912 AGGAGGCCGAGGTGTGCGAACGCTCTGCGAGGAGCTGGGAAACATGGTCCAGGAGCTC 971

Db 781 AGGAGGCCGAGGTGTGCGAACGCTCTGCGAGGAGCTGGGAAACATGGTCCAGGAGCTC 840

QY 972 TCGGGGCTCCAGCTCCTCGTGAAACAGCTCAGCGAGAACCTCAAGAGAGTGTGGAATGAT 1031

Db 841 TCGGGGCTCCAGCTCCTCGTGAAACAGCTCAGCGAGAACCTCAAGAGAGTGTGGAATGAT 900

QY 1032 AACCAAGTTTCTCTGGGAGCTCATTGGTGGCCCTCCTAAGACAAGAAACATGTCAAGTTGC 1091

Db 901 AACCAAGTTTCTCTGGGAGCTCATTGGTGGCCCTCCTAAGACAAGAAACATGTCAAGTTGC 960

QY 1092 TGGCAGGATGGCGGTTCTTTGCGGAAATGAAACGTTGGGTGGTGGACAGCTGCACACAG 1151

Db 961 TGGCAGGATGGCGGTTCTTTTGGGAAATGAAACGTTGGGTGGTGGACAGCTGCACACAG 1020

QY 1152 TGTACTCTCAAGAAATTTTAAACCATTTTGCCACCAATCACTGCCGCCCTGCACACCTGC 1211

Db 1021 TGTACTCTCAAGAAATTTTAAACCATTTTGCCACCAATCACTGCCGCCCTGCACACCTGC 1080

QY 1212 GCCAGTCCATCTCTTTGTGAAGCGAATGTGCGCTTCTCGCTCCCTCCACTCGGTGACCGT 1271

Db 1081 GCCAGTCCATCTCTTTGTGAAGCGAATGTGCGCTTCTCGCTCCCTCCACTCGGTGACCGT 1140

QY 1272 GAGGAGGCTGCTCCGTGGGCGAGAGTGACCCAGTGCTCCGTGACCTGTGGTCTCTGG 1331

Db 1141 GAGGAGGCTGCTCCGTGGGCGAGAGTGACCCAGTGCTCCGTGACCTGTGGTCTCTGG 1200

QY 1332 ACCCAGCAGAGAGGCGGCTCTGTGACGTCAACAGCAACACCTGCTGGGCGCTTCCATC 1391

Db 1201 ACCCAGCAGAGAGGCGGCTCTGTGACGTCAACAGCAACACCTGCTGGGCGCTTCCATC 1260

QY 1392 CAGACACGGCTTGCACTGAGCAGAGTGACACCCGCAATCCCGCAGAGCGCGCTGG 1451

Db 1261 CAGACACGGCTTGCACTGAGCAGAGTGACACCCGCAATCCCGCAGAGCGCGCTGG 1320

QY 1452 AGCCACTGTGTCACTTGTGTTTCTGTGACCTGTGAGTGGCAATATCACACGC 1511

Db 1321 AGCCACTGTGTCACTTGTGTTTCTGTGACCTGTGAGTGGCAATATCACACGC 1380

QY 1512 ATCCGCTCTGCAACTCCCGAGTCCCGCAGTGGGGGCAAGAAATGCAAGGGAGTGGC 1571

Db 1381 ATCCGCTCTGCAACTCCCGAGTCCCGCAGTGGGGGCAAGAAATGCAAGGGAGTGGC 1440

QY 1572 CCGGAGACCAAAAGCTGCGCAGGCGCGCCCAATCGATGCGCTGGAGCGCTTGG 1631

Db 1441 CCGGAGACCAAAAGCTGCGCAGGCGCGCCCAATCGATGCGCTGGAGCGCTTGG 1500

QY 1632 TCCCGTGTGGGCTGCACTGTCACTGTGCGGTGGGATCCCGGAGCGCACCGGGTC 1691

Db 1501 TCCCGTGTGGGCTGCACTGTCACTGTGCGGTGGGATCCCGGAGCGCACCGGGTC 1560

QY 1692 TGCAACAGCCCTCAGACCTCAGTACGAGGGAAGGCGCTCGTGGGGGATGTGAGGAGCGT 1751

Db 1561 TGCAACAGCCCTCAGACCTCAGTACGAGGGAAGGCGCTCGTGGGGGATGTGAGGAGCGT 1620

QY 1752 CAGATGTGCAACAGAGAGAGCTGCCCGTGGATGGCTTTTATCCAAACCCCTGCTCCCG 1811

Db 1621 CAGATGTGCAACAGAGAGAGCTGCCCGTGGATGGCTTTTATCCAAACCCCTGCTCCCG 1680

QY 1812 GGAGCCAGTGCAGCAGCTTCCCGATGGGTCTCTGATCGCGCTCTGCTGCGCTGTGGGC 1871

Db 1681 GGAGCCAGTGCAGCAGCTTCCCGATGGGTCTCTGATCGCGCTCTGCTGCGCTGTGGGC 1740

QY 1872 TTCTTGGGCAATGGCAACCACTGTGAGACCTTGACAGTGTGCCCTGGTCCCGGACATC 1931

Db 1741 TTCTTGGGCAATGGCAACCACTGTGAGACCTTGACAGTGTGCCCTGGTCCCGGACATC 1800

QY 1932 TGCTTCTCCACCAAGAGGTCCTCGCTGTGTCAACACTCAGCCTTGGCTTCCACTGCTCG 1991

Db 1801 TGCTTCTCCACCAAGAGGTCCTCGCTGTGTCAACACTCAGCCTTGGCTTCCACTGCTCG 1860

QY 1992 CCTTCCCGCCCGGATACAGAGGAAACAGCCGCTCGGGGTGGGCTGGAGAGGACGACG 2051

Db 1861 CCTTCCCGCCCGGATACAGAGGAAACAGCCGCTCGGGGTGGGCTGGAGAGGACGACG 1920

2052 ACAGAAAGCAAGTGTGTGAGCCCGAAACCCATGCAAGCAGACAGACACAACTGCCAC 2111
2121 ACAGAAAGCAAGTGTGTGAGCCCGAAACCCATGCAAGCAGACAGACACAACTGCCAC 1980
2112 AAGCACGGGAGTGCATCTACCTGGGCCCACTTTCAGCGACCCCATGATGACAGTCCGAGTGC 2171
1581 AAGCACGGGAGTGCATCTACCTGGGCCCACTTTCAGCGACCCCATGATGACAGTCCGAGTGC 2040
2172 CAGACAGCTAGCGGGGCGACGGGCTCATCTCGGGGAGGACTCGGACCTGGACGGCTGG 2231
2041 CAGACAGCTAGCGGGGCGACGGGCTCATCTCGGGGAGGACTCGGACCTGGACGGCTGG 2100
2232 CCCAACTCAATCTGTCTGTGCGCCCAACCAACGACCTTACCTGCACTCAAGGATAACTGC 2291
2101 CCCAACTCAATCTGTCTGTGCGCCCAACCAACGACCTTACCTGCACTCAAGGATAACTGC 2160
2292 CCCCACTGCCAAATCTGGGCGAGGAAGACTTTTGAACGAGCGGATTTGGCGATGCTGT 2351
2161 CCCCACTGCCAAATCTGGGCGAGGAAGACTTTTGAACGAGCGGATTTGGCGATGCTGT 2220
2352 GATGATGACGATGACAAATGACCGTGTGACCGATGAGGAGGACAACTGCGAGCTCCTCTTC 2411
2221 GATGATGACGATGACAAATGACCGTGTGACCGATGAGGAGGACAACTGCGAGCTCCTCTTC 2280
2412 AATCCCGCCAGGCTGACTATGACAAAGGATGAGTTGGGGAACCGCTGTGACAACTGCCCT 2471
2281 AATCCCGCCAGGCTGACTATGACAAAGGATGAGTTGGGGAACCGCTGTGACAACTGCCCT 2340
2472 TACGTGACAACTCCCGCCAGATCGACACAGACAAACAAATGGAGAGGCTGACGCTGCTCC 2531
2341 TACGTGACAACTCCCGCCAGATCGACACAGACAAACAAATGGAGAGGCTGACGCTGCTCC 2400
2532 GTGACATTTGATGGGAGGATGCTTCTCAATGAACGAGACAAATTTGCCCTACGTTCAAC 2591
2401 GTGACATTTGATGGGAGGATGCTTCTCAATGAACGAGACAAATTTGCCCTACGTTCAAC 2460
2592 ACTGACAGAGGAGACAGGATGGTGAACGTTGGGGATCACTGTGACAACTGCCCTG 2651
2461 ACTGACAGAGGAGACAGGATGGTGAACGTTGGGGATCACTGTGACAACTGCCCTG 2520
2652 GTGACAACTCCCGCCAGATCGACACAGGACCAATGACCTTTGGGAGACCACTGTGACAA 2711
2521 GTGACAACTCCCGCCAGATCGACACAGGACCAATGACCTTTGGGAGACCACTGTGACAA 2580
2712 AACGAGGACATAGATGACAGCGGCCACCAAGAACACAGGACCACTGCCCTACATCTCC 2771
2581 AACGAGGACATAGATGACAGCGGCCACCAAGAACACAGGACCACTGCCCTACATCTCC 2640
2772 AACGCAACCTGAGCTGACCATGACAGAGCGGCCAGCGCCGTGTGACCTCTGATGAT 2831
2641 AACGCAACCTGAGCTGACCATGACAGAGCGGCCAGCGCCGTGTGACCTCTGATGAT 2700
2832 GACAAAGATGGGCTCCCGATGACAGGAGCACTGCGGGTTGTGTTCAACCCAGACAG 2891
2701 GACAAAGATGGGCTCCCGATGACAGGAGCACTGCGGGTTGTGTTCAACCCAGACAG 2760
2892 GAGGACTTGGACGGTGTGACAGCGGCTGATATTGTAAAGATGATTTGACAAATGACAA 2951
2761 GAGGACTTGGACGGTGTGACAGCGGCTGATATTGTAAAGATGATTTGACAAATGACAA 2820
2952 ATCCAGATTTGATGATGTGTCTGAAAACAAATGCCATGATGAGACAGACTTCAGG 3011
2821 ATCCAGATTTGATGATGTGTCTGAAAACAAATGCCATGATGAGACAGACTTCAGG 2880
3012 AACTTCCAGATGCTCCCTTGGATCCCAAGGAGCAACCAATTTGATCCCACTGGGTC 3071
2881 AACTTCCAGATGCTCCCTTGGATCCCAAGGAGCAACCAATTTGATCCCACTGGGTC 2940
3072 ATTGCCATCAAGGCAAGGAGCTGTTTCAGACAGCACTCGGACCCCGCATGCTGTA 3131
2941 ATTGCCATCAAGGCAAGGAGCTGTTTCAGACAGCACTCGGACCCCGCATGCTGTA 3000
3132 GGTTTTGACGAGTTTGGGTCTGTGGACTTCAGTGGGCACATTTCTACGTAACACTGACCGG 3191

3001 GGTTTTGACGAGTTTGGGTCTGTGGACTTCAGTGGGCACATTTCTACGTAACACTGACCGG 3060
3192 GACACGACTATGCCGGTTCTGTTGGTTACCACTAAGCAGCGCTTCTATGTGGTG 3251
3061 GACACGACTATGCCGGTTCTGTTGGTTACCACTAAGCAGCGCTTCTATGTGGTG 3120
3252 ATGTGGAGAGGTTGACCGCAGACCTTACTGGGAGGACCCACCGGGCTTATGGCTAC 3311
3121 ATGTGGAGAGGTTGACCGCAGACCTTACTGGGAGGACCCACCGGGCTTATGGCTAC 3180
3312 TCCGCGGTGTCCCTCAAGGTGTGTAACCTCCACCGGGGACCGGCGAGCACTGAGGAAC 3371
3181 TCCGCGGTGTCCCTCAAGGTGTGTAACCTCCACCGGGGACCGGCGAGCACTGAGGAAC 3240
3372 GCGTGTGGCACACGGGGAACACGCGGGGAGGTGCGAACCTTATGGACGACCCCGG 3431
3241 GCGTGTGGCACACGGGGAACACGCGGGGAGGTGCGAACCTTATGGACGACCCCGG 3300
3432 AACATTGGCTGGAAGGACTTACACGCGCTTATAGTGGACCTTGACTCACAGGCCCAAGACT 3491
3301 AACATTGGCTGGAAGGACTTACACGCGCTTATAGTGGACCTTGACTCACAGGCCCAAGACT 3360
3492 GGCTACATCAG 3502
3361 GGCTACATCAG 3371

RESULT 2
AY398792 3371 bp DNA linear GSS 15-DEC-2003
LOCUS Mus musculus THBS2 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY398792
VERSION AY398792.1 GI:39754784
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 3371)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
Gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3371)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
source location/Qualifiers
1..3371
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
1..3371
/gene="THBS2"
/locus_tag="HMC0008"

gene
Query Match 43.6%; Score 2486.2; DB 29; Length 3371;
Best Local Similarity 83.6%; Pred. NO. 0;
Matches 2818; Conservative 0; Mismatches 553; Indels 0; Gaps 0;

Qy	132	ATG	G	T	C	T	G	T	G	A	G	G	C	T	G	T	C	T	G	G	T	G	T	G	G	C	C	A	G	C	T	G	G	T	C	A	C		191
Db	1	A	T	G	C	T	C	T	G	G	C	A	T	G	G	C	T	G	G	C	A	T	G	G	C	C	A	G	A	G	C	T	T	G	C	T	G	C	60
Qy	192	C	A	G	C	A	A	A	G	A	C	A	G	C	A	C	T	T	C	G	A	C	T	T	C	A	G	A	T	C	A	G	A	A	C	A	C	251	
Db	61	C	A	C	G	T	C	A	A	G	G	A	C	A	C	T	T	C	A	T	T	G	A	C	T	T	C	A	G	A	C	A	T	A	A	C	G	120	
Qy	252	G	G	C	C	A	A	G	C	A	G	T	T	C	G	G	G	C	C	G	A	C	C	C	G	G	C	T	T	A	C	C	G	T	T	C	G	311	
Db	121	G	G	T	G	C	A	A	G	C	A	G	T	T	C	G	A	G	G	C	T	G	A	C	C	C	G	G	T	G	C	C	G	T	T	G	A	180	
Qy	312	G	A	C	T	A	C	T	C	C	C	A	C	C	G	T	G	A	C	G	C	A	G	A	T	A	C	T	C	A	G	A	G	A	T	C	371		
Db	181	G	A	C	T	A	C	T	C	C	C	C	A	G	T	G	A	A	C	A	G	A	T	G	T	C	A	A	G	A	T	T	G	T	C	A	240		
Qy	372	A	A	G	A	G	G	C	T	T	C	T	C	A	C	G	G	C	C	A	G	T	C	A	A	G	A	C	G	C	A	G	C	A	G	C	431		
Db	241	A	A	G	A	G	G	C	T	T	C	T	C	A	C	A	G	C	C	A	A	T	A	G	A	C	G	A	G	C	C	A	G	T	C	G	300		
Qy	432	T	T	G	C	T	C	T	G	A	G	G	C	C	C	G	T	C	T	C	C	A	G	A	G	T	T	C	G	A	T	T	C	C	A	C	491		
Db	301	C	T	G	T	T	G	A	A	G	C	C	C	G	C	A	C	C	T	C	C	A	G	A	G	T	T	G	A	T	T	G	T	T	G	T	360		
Qy	492	C	G	G	A	C	G	C	T	G	A	T	C	T	A	C	T	A	G	A	T	T	G	A	C	G	C	A	C	C	G	C	A	T	G	T	551		
Db	361	G	G	G	A	C	A	T	T	G	A	C	T	C	A	A	C	T	A	C	T	G	G	T	A	G	A	C	A	A	T	A	G	A	C	A	420		
Qy	552	G	A	C	G	T	C	G	C	T	G	C	T	A	C	G	A	T	G	A	A	A	C	G	T	C	A	C	C	G	T	C	A	C	G	A	611		
Db	421	G	A	T	G	T	G	G	G	C	T	G	C	T	C	C	A	T	G	A	A	A	T	G	T	A	C	T	G	T	C	A	G	T	G	C	480		
Qy	612	T	A	C	A	G	T	T	C	A	G	T	C	S	C	G	A	C	T	C	A	G	A	G	T	T	C	G	T	C	T	C	G	A	C	G	671		
Db	481	T	A	T	A	G	C	T	G	T	T	G	G	C	T	C	G	A	T	T	C	T	T	A	T	C	G	A	C	A	G	T	C	A	C	A	540		
Qy	672	G	A	C	A	C	T	C	A	G	C	G	A	A	A	G	A	G	C	G	A	T	G	T	A	C	G	T	G	C	C	A	A	G	G	T	731		
Db	541	G	A	G	C	A	G	T	A	G	A	T	A	G	A	C	A	G	A	G	C	A	G	A	T	G	T	A	C	T	G	G	C	A	A				

[illegible]

2161	CCCAAAGCTGCCAAATTC	CGCGCGAGGAGGATTTT	TGATTAAGGATCGGAATCGGAGATGCTTGT	2222
2352	GATGATGACGATGACAAATGA	CGGTGTGACCGATGAGAAGGACAACT	GCCAGCTCCTCTTTC	2411
2221	GACGAGCAGCATGACAAATGAC	CGGTGTGAGCGATGAGAAGGACAAAT	TTCGCACTTCTCTTC	2280
2412	AATCCCGCCGAGCTGACTAT	GACAGGATGAGGTTGGGACCGCTGTGACAACTGCCCT	2471	
2281	AATCCCGCTCAATTAGACT	ATGACAAAGGATGAGTTGGAGACCGCTGTGACAACTGCCCC	2340	
2472	TACGTGACAAACCCCTG	CCAGATCGACACAGACAAATGAGAGGGGTGACGCTGCTCC	2531	
2341	TATGTGCACAACCCAGCAC	AGATCGACACAGACAAATGCGGAGGGGATGCGCTGCTCT	2400	
2532	GTGGACATTTGATGGGAC	CGATGCTTCTCAATGAACGAGACAAATTTGCTCCCTACGCTCTACAAC	2591	
2401	GTGGACATTCACGGAGAC	GATGTTTCAATGAGCGGAGACAAATTTGCTCATATGTCTACAC	2460	
2592	ACTGACAGAGGAGACACG	GATGGTGTGGGGATCACTGTGACAACTGCCCCCTG	2655	
2461	ACTGACAGAGAGATACG	GATGGTGTGGGTGACCACTGTGACAAATTTGCTCTCTG	2520	
2652	GTGCACAAACCCCTGAC	CAGACCGGACAAATGACCTTGTGGGACCACTGTGACAACT	2711	
2521	ATGCACAAACCCAGATC	AGATCAGGACAAATGATCTCGTTGGAGACCACTGTGACAACT	2580	
2712	AACGAGACATAGATGAC	AGCGCCACAGAACACGAGGACAACTTGCCCTCATCATCTCC	2771	
2581	AATGAGGACATAGATG	ATGACGGCCACAGAACACGAGGACAACTTGCCCATACATCTCC	2640	
2772	AACGCCAAACGAGCTG	ACCATGACAGAGACGGCCAGGGCGACCGCTGTGACCTGATGAT	2831	
2641	AACTCCAAACGAGCTG	ACCATGACACGCGCAAGGCGATGCTCGCTGCGACTCTGATGAT	2700	
2832	GACAAAGATGGCGTCC	CGATGACAGGACAACTGCGCGGTGTGTTCAAACCGAGACCCAG	2891	
2701	GACAATGATGGTGTTC	AGATGACAGGACAACTGTGGGTGTGTTCAAACCGAGACCCAG	2760	
2892	GAGGACTTGACCGGTG	ATGACGGGGTGATTTGTAAAGATGATTTGACAAATGACAACT	2951	
2761	GAAGACTCGGACGGT	GACGCGCGAGGTGACATTTGTAAAGATGACTTTGACAAATGATAT	2820	
2952	ATCCAGATATTGATG	ATGTGCTGAAAAAATATGCCATGATGATGAGACAGACTTCAGG	3011	
2821	GTCCAGATATTGATG	ATGTGCTGAAAAAATATGCCATGATGATGAGACAGACTTCAGA	2880	
3012	AACTTCCAGATGTGCC	CTTGGATCCCAAAAGGAGACCAACCAAAATTGATCCCAACTGGGTC	3071	
2881	AACTTCCAGATGTGCC	CTTGGATCCCAAGGAGACCAACCAAAATTGATCCCAACTGGGA	2940	
3072	ATTCGCCATCAAGG	CAAGGAGCTGGTTCAGACAGCGCACTCGGACCCCGGATCGCTGTA	3131	
2941	ATTTCGTCAACAGG	CAAGGAGCTGGTTCAGACAGCAAACTCAGACCCCTGGCATCGCTGTA	3000	
3132	GGTTTTGACAGATTT	TGGGCTGTGGACTTCAGTGGGACATTTCTACGTAAACACTGACCCGG	3191	
3001	GGTTTTGACAGATTT	TGGGCTGTGGACTTCAGTGGGACTTTCTATGTCAACACTGACCCGG	3060	
3192	GACGACGACTATG	CGCGCTTCTTTGGTTTACAGTCAAGGAGCCGCTTCTATGTGGTG	3251	
3061	GATGATGACTACG	CTTGGCTTCTTTCAGTCAAGGAGCCGCTTCTATGTGGTG	3120	
3252	ATGTGAAGCAGGTG	ACGAGACTCTGGGAGGAGCCAGCGCCACCGCGGCTATGGCTAC	3311	
3121	ATGTGAAGCAGGTG	ACGAGACTCTGGGAGGAGCAAGCCGCTTACCGGCTAC	3180	
3312	TCGGCGGTGCTC	TAAGGTGGTGAACCTCCACCGGGGACGGCGGACCTTGAGGAAC	3371	
3181	TCGTGTGTCACTCA	AAAGTGGTAACTCCACGACTGGTACTGGCGAGACCTTGAGGAAT	3240	
3372	GCCTGTGGCAC	CGGGGAACGCGCGGGCAGGTGCGAACTTTATGGCACGCCACCG	3431	
3241	GCCTGTGGCAC	CGGGGAACAGAAAGGCGAGGTCCGGACTCTATGGCATGACCCCAA	3300	

QY	3432	AACTTGGCTGGAAGACTACACGGCCCTATAGTGGGACCTGACTCACAGGGCCCAAGACT	3439
Db	3301	AACTTGGCTGGAAGACTACACGGCCCTATAGTGGGACCTGACTCACAGGGCCCTAAGACA	3360
QY	3492	GGCTACATCAG	3502
Db	3361	GGCTACATGAG	3371
RESULT 3			
LOCUS	AY398791		
DEFINITION	Pan troglodytes THBS2 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		
ACCESSION	AY398791		
VERSION	AY398791.1	GI:39754783	
KEYWORDS	GSS.		
SOURCE	Pan troglodytes (chimpanzee)		
ORGANISM	Pan troglodytes		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.		
AUTHORS	1. (bases 1 to 3273)		
	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.		
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science	302 (5652), 1960-1963 (2003)	
PUBMED	14671302		
AUTHORS	2. (bases 1 to 3273)		
	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-NOV-2003)	Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA	
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.		
FEATURES	Location/Qualifiers		
source	1..3273		
gene	/organism="Pan troglodytes"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9598"		
	<1..>3273		
	/gene="THBS2"		
	/locus_tag="HCM0008"		
ORIGIN			
Query Match	42.3%;	Score 2408.8;	DB 29; Length 3273;
Best Local Similarity	73.8%;	Pred. No. 0;	
Matches 2416;	Conservative 0;	Mismatches 857;	Indels 0; Gaps 0;
QY	132	ATGGTCTGAGGCTGCTCTGCTGGCTCTGTTGGGCTGCGGCAGCACGCAAGCTGCTCAC	191
Db	1	ATGGTCTGAGGCTGCTCTGCTGGCTCTGTTGGGCTGCGGCAGCACGCAAGNNNNNN	60
QY	192	CAGGCAAAAGACACGACCTTTTCAGTATCAGCAACATCAACCGCAAGCACTT	251
Db	61	NN	120
QY	252	GGCGCCAAAGAGTTCGCGGGCCGACCCCGCGCTTACCGCTTCGTCGCTTT	311
Db	121	NN	180
QY	312	GACTACATCCACCGGTGAACGAGATGACCTCAGCAAGATCACCAAGATCATCGGCAG	371
Db	181	NN	240
QY	372	AAGGAGGGCTTCTTCCTCAGCGGCCAGCTCAAGCAGGACGCGCAAGTCCAGGGCAGCTG	431

/tissue_type="cortex"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
217..3733
misc_feature
/note="THROMBOSPONDIN 1 PRECURSOR homolog [Mus musculus]
(SWISSPROT|P35441, evidence: PASTV, 99.7%ID, 100%length,
match=3514)
putative"

ORIGIN

Query Match 26.5%; Score 1510.8; DB 11; Length 4109;
Best Local Similarity 66.4%; Pred. No. 4.5e-216;
Matches 2258; Conservative 0; Mismatches 1112; Indels 30; Gaps 5;
QY 270 GGGCCGACCCCGCGCTTACCGCTTCCGCTTTCGCGCTTGTGACTACATCCCAACCCGCTG 329
DB 367 GGCAGATCTATCCAGCCCGCTTCCGATTTAGAAATGCAACCTGATCCCGCTGTG 426
QY 330 AACGAGATGACCTCAGCAAGATCAACAGATCATCGGCAGAGAGGGTTCTTCTC 389
DB 427 CCGGATGACAAAGTTCCAAAGACCTACTGGACGCTGTGTGGCCGACAAAGGCTTCATCTC 486
QY 390 ACGGCCACCTCAAGCAG-GACGGCAAGTCCAGGGCAGCTGTGGCTCTGGAGGGGCC 448
DB 487 CTGGCTCTTGGAGCAGATGAAAAAACCAGGGGCACTCTCTGCTGTGGAACGGAA 546
QY 449 CGGTCTCTCCAGAGCAGTTCCAGATTCGTCTCAACGGCCCGCGACACGCTGGATCT 508
DB 547 AGAAAAACCTGGCCAGATCTTCACTAGTGTGGTCTCCACCGCAAGCTGGCACCTTGGACCT 606
QY 509 CACCTACTGATGAGCGCACCGGATGTGCTCTCTGGAGGACGTTCGSCCTGGCTGA 568
DB 607 GAGCTGAGCTGCGCAGGAGCAACAGTGTGTCTAGTGGAGGAGCTCTCTTGGCCAC 666
QY 569 CTCGAGTGGAGAACCTCACCTGCAAGTGGTGGCGAGACCTACAGCTTGCACGTGGG 628
DB 667 TGGCCAGTGGAGAGCATCACGCTGTGTTCAAGAGGACCGGGCTCAACTCTACATAGA 726
QY 629 CTGCGACCTCATAGACAGCTTCCTCTGGACGAGCCCTTCTACGACACCTGCGAGGGGA 688
DB 727 CTGTGATAGATGGAGAGCGGAGCTGGATGTATACCATCA-GAGCATCTTACCAGGG 785
QY 689 AAAGAGCCGGATGTACGTGGCCAAAGGCTCTGCGCAGAGAGTCTC-----ACTTCAGGG 741
DB 786 ATCTGCCAGCGTGGCCAGGCTCCAGTTGCAAGGAGAGATGTCATGACAAATTTTCAGG 845
QY 742 GTTGTCTTCCAGACGCTCCACTAGTGTGTTGAAGTCTGTGGAGATATTCTAAGCAGA 801
DB 846 GGGTGTCTGAGAAATGTAGGTTTGTCTTGGAAACACCCCAAGACATCTCAGGAACA 905
QY 802 AGGGTTGCCAGAACGCGCAGGAGCTGAGATCAACGCCATCAGTGAAGAACACAGAGCGC 861
DB 906 AAGGCTGC---TCCAGCTCAGCTACCAACGCTCTTCTTACCCCTTGACAAACGCTGTGA 962
QY 862 TGGCCTGGTTCGCGATGTCCACACCGAGTACGTGGGCCCGCCAGCTCAGAGAGGAGGCCG 921
DB 963 ACGGTTCCAGCCCTGTATCCGACCAACTATATCGGCCCAACAAAGAGACCTCCAAG 1022
QY 922 AGGTGTGCGAACGCTCGTGGAGAGCTGGGAACATGGTCCAGAGCTCTCGGGGCTCC 981
DB 1023 CTATCTGTGGCTCTCTCTGTGTGATGAACCTATCCAGCATGGTCTCTGAACTGAAGGCGCTGC 1082
QY 982 AGTCTCTGTGAAACAGCTCAGCGAGAACCTCAAGAGAGTGTGCAATGATAACCAAGTTTC 1041
DB 1083 GCACCATCTGACCACTCTCGAGGACAGCATCCGAAAGTGCAGGAGAGAACAGAGAGC 1142
QY 1042 TCTGGAGGCTCATTTGTGGCCCTCTCTAAGACAAAGAACATGTCACTAGTTTGGCGAGGATG 1101
DB 1143 TGTCTAGTGAAGTGAAGCGGCTTCCCTC-----TGCTTTTCACAATG 1184
QY 1102 GCGGTTCTTTGGGAAATGAAACGTTGGTGGAGAGCTGCACACGCTGTACTCTCA 1161
DB 1185 GAGTCCAGTACAAACACACGAGGAGTGGACTGTAGACAGTTGACAGAGTGTCACTGCC 1244

QY 1162 AGAAATTTAAACCATTTGCCACCAATCACTGCGCCCGCTGCAACCTGCGCCAGTCCAT 1221
DB 1245 AGAACTCGTGTACCATCTGCAAAAAGGTGCTCTCCCATCATGCTCCCTGCTCCAAACGCA 1304
QY 1222 CTTTGTGGAAGCGCAATGCTGCTCTTCTGCTCTCCACTCGGTGGAGCGGTGAGGAGGCT 1281
DB 1305 CAGTTCTCTGATGGTGAATGCTGCTCCACCGTGTGCGCCACGAGCTCTGCTGACGATGCT 1364
QY 1282 GGTCTCGTGGCAGAGTGGACCCAGTGTCTCGTGTGCTGTGGCTTGGGACCCAGCAGA 1341
DB 1365 GGTCTCCCTGGTGTGAGTGGACCTCTGCTCTGCCCATGTGGCAATGGAATTCAGCAAC 1424
QY 1342 GAGGCGCGTCTGTGACGTCAACAGCAACACCTGCTTGGGCGCTCTCATCCAGACACGGG 1401
DB 1425 GTGGTCTGTTCTGTGACAGCTCAACACAGATGCGAGGGCTCTTCGTGTACAGACAAG 1484
QY 1402 CTTGAGTCTGAGCAAGTGTGACACCGCATCCGGAGGACCGGCTGGAGCCACTGGT 1461
DB 1485 CTTGCCACATTCAGGAGTGTGACAAAAGATTTAAACAGGATGGTGGCTGGAGTCACTGGT 1544
QY 1462 CACCTTGGTCTTCTCATGCTCTGTGACCTGTGGAGTTGGCAATATCACACGCTATCCGCTCT 1521
DB 1545 CTCCATGGTCTGCTCTGCTGTGACCTGTGCTGTGATCACAAGGATCCGGCTCT 1604
QY 1522 GCAACTCCCCAGTGGCCAGATGGGGGCAAGAAATTCGAAAGGAGTGGCCGGGAGACA 1581
DB 1605 GCAACTCCCCAGCCCGCCAGATGAACGGGAAGCCCTGTGAAGGTGAAGCCCGGGAGACA 1664
QY 1582 AAGCCTGCCAGGGCGCCCATGATGATGCGCTGAGAGCCCTGTGCTCCCGTGGT 1641
DB 1665 AAGCCTGCAAGAAAGACGCTGCTCCCAATTAATGGAGGCTGGGGTCCCTGGTCAACATGG 1724
QY 1642 CGGCTGCTCATCTGCTGCTGCTGGATCCGGAGCGCACCGGCTCTGCAACAGCC 1701
DB 1725 ACATCTGCTCTGTGCTGCTGAGGAGGAGTGCAGAGACGCGAGCTCTGTAAACAAC 1784
QY 1702 CTGAGGCTCAGTACGAGGAGGAGCGCTGCTGCGGGATGTGCAGAGAGCGGTGAGATGTGA 1761
DB 1785 CCACACCCAGTTTGGAGGCAAGACATGCTGTGTCGATGTGACAGAAAAATCAAGTTTGA 1844
QY 1762 ACAGAGGAGTGTCCCGTGGTGGATGGCTGTATTCACACCTGCTTCCCGGAGCGGCT 1821
DB 1845 ACAAGCAGGACTGCCCAATTTGATGATGCTGCTGCAATCCCTGCTGTGCTGGTGGCAAGT 1904
QY 1822 GCAGCAGCTTCCCGATGGTCTGCTGCTCATGCGGCTCTGCTGCTGCTGCTGCTGCTGCTG 1881
DB 1905 GTACTAGCTACCTGATGGTAGCTGGAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1964
QY 1882 ATGGCAACCACTGTGAGGACCTGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1941
DB 1965 ATGGCATCCAGTGCAGGAGAGCTGATGAGTGCAGGAGAGTGGCTGATGCTGCTGCTGCTG 2024
QY 1942 CCAGCAAGGTGCTGCTGCTGCTCAACACTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2001
DB 2025 ACAACGAGAACATCGTGCAGAACACACAGATCTCTGCTCAACTGCTGCTGCTGCTGCTGCTG 2084
QY 2002 CCGGATACAGAGGAACACAGCCGCTCGGGGTGCGCTGGAAGAGCAGCAGGAGGAGG 2061
DB 2085 CACGATTCTGCTGCTCAGAGCCCTTCCGCGAGGTGTCGAACATGCCATGGCCAAACAAAC 2144
QY 2062 AAGTGTGTGAGCCCGCAAAACCCCATGCAAGGACAAAGACACAACTGCCACAGCAGCGG 2121
DB 2145 AGTGTGTCAACCGCAACCCCTGCAAGGAGCGGACGATGATGCAACAAAGAACGCTA 2204
QY 2122 AGTGTACTTCTGCGGCTTCTGAGCGACCCCATGTACAAGTGGAGTGGAGGAGGAGGCT 2181
DB 2205 AGTGAACACTACCTGGGTCACTACAGTACCCCATGTACCGCTGTGAGTGCAGAGCCGGCT 2264
QY 2182 ACGCGGCGCAGCGGCTCATCTGCGGGGAGGACTCGGACCTCGGAGCGGCTGGCCCAACCTCA 2241
DB 2265 ATGCAGGCAATGCAATCATCTGCGGAGAGGACACAGACCTGGAAGGCTGGGCTGCTAATGAA 2324

2242 ATCTGGTCTGCGCCACCAACCGCACCCTACCTGATCAAGGATTAAGTCCCGCCATCTGC 2301
2325 ACCTGGTGTGTGCGCCAAACGCACTACCTGCAAAAAGGACAACTGCCCACTTC 2384
2302 CAATTTGGCGAGGAAGACTTTGCAAGAGCGGATTCGGCATGCTGTGATGATGACG 2361
2385 CCACTCGGGGAGGAAGACTATGCAAGAGCGGATTCGGCATGCTGTGATGATGACG 2444
2362 ATGACAATGACGGTGTGACCGATGAGAGGACAACTGCCAGCTCCCTTCAATCCCGCC 2421
2445 ATGACAAGCAAGATCCCGATGACAGGACAACTGTCTCATTTCAATCAACCCAGCCC 2504
2422 AGCTGACTATGACAGGATGAGTTGGGGACCGCTGTGACAACTCCCTTACGTGCACA 2481
2505 AGTATGACTATGACAGATGATGTGGGAGACCGCTGTGACAACTGCCCTTCAACCA 2564
2482 ACCCTGCCAGATGCACACAGACAACTGAGAGGTCAGCGCTCCTCGTGACATTCG 2541
2565 ACCCTGACCAAGACAGACAAAGAACCGGGAGGGCGATGCTGTGTGGATGTCG 2624
2542 ATGGGACGATGTCTTCAATGACAGACAAATGTCCCTACGCTCTACAACTGACGACA 2601
2625 ATGGAGATGGAATCTCAATGAACGAGACAACTGCCAGTACGTTTACAACTGACCA 2684
2602 GGGACAGGATGGTGCAGGTGTGGGGATCACTGTGACAACTGCCCTGTGACAACT 2661
2685 GGGACAGGATGATGGGTGGGATGATGATGACAACTGCCCTGTGACAACT 2744
2662 CTGACCAAGACCGCTGACCAATGACTTGTGGGACCAAGTGTGACAACTGACGAGCA 2721
2745 CAGACAGCTGACTCTGACTCAGACCTCATAGGGGACACTTGTGACAACTGACGACA 2804
2722 TAGATGACGCGGACACAGACAACTGACGACAACTGCGCTTACATCTCCAGCGCACT 2781
2805 TCGATGAGGATGGCATCAGAACAACTGTGACAACTGTCCCTATGCGCTTAAATGCCA 2864
2782 AGGCTGACCATGACAGACAGCGGACGAGCGCTGTGACCGCTGATGATGACAACTG 2841
2865 AGCGGACCATGATAAGATGCAAGGAGATGCTGTGACCATGACGATGACAACTGACG 2924
2842 CGCTCCCGATGACAGGACAACTGCGGCTGTGTGTTCAACCCAGACGAGGAGCTTGG 2901
2925 GCATCCCTGATGACAGACAACTGCGGCTGTGTGCGCAATCTGACCAAGAGGACTCTG 2984
2902 ACGGTGATGGACGGGTGATATTTGTAAGATGATTTTGACATGATGACAACTCCAGATA 2961
2985 ATGTGATGGCGAGTGACGCTGCAAGACGACTTTGACCATGACATGTCGCGAGATA 3044
2962 TTGATGATGTGTCTGAAACAAATGCCATCAGTGAGACAGACTTCAGGAACTTCGAGA 3021
3045 TTGATGACATCTGTCTGAGAAATTTTGACATCAGTGAAACCGGATTTCCGACAAATTC 3104
3022 TGGTCCCTTGGATCCCAAGGAGCACCCAAATTTGATCCCACTGGTTCATTCGCCATC 3081
3105 TGATTCCTTAGATCCCAAGGAACTCCCAAAATGACCTTAAGCTTGGTGTGCGCATC 3164
3082 AAGCAAGGAGCTGTTGACAGACGCAACTCGGACCCCGCATCGCTGTAGGTTTGAAG 3141
3165 AGGCAAGAACTTGTCCAGACTGTAAACTGTGACCCCTGACCTGTGTAGGTTATGATG 3224
3142 AGTTTGGTCTGTGACTTCACTGACATTTCTACCTAACTGACCGGAGCAGCACT 3201
3225 AGTTTAACTGTGAGATTCACGGGTACCTTCTTCAACACCCGAAAGATGATGACT 3284
3202 ATGCCGGTCTGTCTTTGGTTACAGTCAAGCAGCGCTTCTATGTGGTGATGTGGAAGC 3261
3285 ACGTGGCTTTGTTTGGCTACCACTGTCAGCAGCGCTTCTACGTGTGTGATGTGGAAC 3344
3262 AGGTGACCGACCTTACTGGGAGGACCCAGCGGGCTATGCTACTCTCGGCTGT 3321
3345 AAGTACCGAGCTTACTTGGGACCAACCCCAAGGGCTCAGGGATCTCAGGCTGT 3404
3322 CCTCAAGTGTGTGAATCTCCACACGGGACGGGCGAGCACTGTGAGAAACGCGCTGTGGC 3381

Db 3405 CTGTAAAGTTGTGAACCTCCACACCGCCCTGGCGAGCACCTGGGAATGCACTGTGGC 3464
Qy 3382 ACACGGGAAACACGCCGGGCGAGGTGCGAACCTTATGGCACGACCCAGGAAACATTTGGCT 3441
Db 3465 ACACAGGAAACACCCCTGGCCAGGTGCGCACCTGTGGCATGACCTCGCCACATCGCT 3524
Qy 3442 GGAAGGACTCACAGCCCTATAGTGGCAGCTGACTCACAGGCCCAAGACTGGCTTACATCA 3501
Db 3525 GGAAGATTTCACTGCATACAGATGGCGTCTCAGCCACAGGCCAAAGACCGTTATATCA 3584
Qy 3502 GAGCTTTAGTGCATGAAGAAACAGGTTCATGGCAGACTCAGGACCTATCTATGACCAAA 3561
Db 3585 GAGTGGTGTATGATGAAGAAAGAAAATCATGGCTGACTCGGAGCCATCTATGACAAA 3644
Qy 3562 CCTACGCTGGCGGGCGCTGCTATTTGTCTTCTCTCAAGAAATGGTCTATTTCTCAG 3621
Db 3645 CCTACGCGCGGTAGACTAGCGCTTGTCTTCTCTCAGGAATGGTGTCTTCTCAG 3704
Qy 3622 ACCTCAAGTACGAATGCAGATATTTAAACAAGATTTCG 3661
Db 3705 ACATGAATACGAGTGTGAGATTCCTATCATCAGCTGC 3744

RESULT 5

AK089315/c
LOCUS AK089315 3530 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus B6-derived CD11 +ve dendritic cells cDNA, RIKEN
full-length enriched library, clone:F730004F21
product:thrombospondin 1, full insert sequence.

ACCESSION AK089315

VERSION AK089315.1 GI:26105200

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

REFERENCE

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

2

```

of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3530)
JOURNAL
REFERENCE
AUTHORS
Adachi,J., Alzawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Murmatsu,M., and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC Building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
FEATURES
source
1. 3530
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:F730004F21"
/db_xref="MGI:2428221"
/db_xref="taxon:10090"
/clone="F730004F21"
/cell_type="36-derived CD11 +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
misc_feature
1..3530
/notes="thrombospondin 1 (MGI:98737, GB|NM_011580,
evidence: BLASTN, 99%, match=2839)"
ORIGIN
Query Match 24.9%; Score 1416.6; DB 11; Length 3530;
Best Local Similarity 69.9%; Pred. No. 5.6e-202;
Matches 1975; Conservative 0; Mismatches 829; Indels 21; Gaps 4;
QY 840 ATCAGTGAGACACAGAGACGCTGGCTGGTCCGATGTCACACCGAGTACCTGGGC 899
DB 3487 ACCCTTTGACAAACACGTTGGTGAAGCGTTCCAGCCCTGCTATCCGACCAACTACATCGGC 3428
QY 900 CCCAGCTCAGAGAGGAGCCGCGAGTGTGCGAACCGCTCGTGGAGAGTGGGAAACATG 959
DB 3427 CACAAACAAAGGACCTCCAGCTATCTGTGGCTCTCTCTGTGATGAATATCCAGCATG 3368
QY 960 GTCCAGGAGTCTCGGGGCTCCAGTCTCTGTGAACCAAGTCTCAGGAGNACCTCAAGAGA 1019
DB 3367 GTCTTGAATGAAGGGCTCGCGACCATCGTGACCACTCTGTGAGGAGAGCATCCGAAA 3308
QY 1020 GTGTGGAATGATACCAAGTTTCTCTGGGAGTCTATTGGTGGCCCTCCCTAAGACAAGAAC 1079
DB 3307 GTGACGAGAGAGACAGAGCTGTCTAGTGAGCTGAAGCGGCTCCCTCC----- 3257
QY 1080 ATGTCAGCTTCTGGCAGGATGGCGGTTCTTTCGGGAAATGAACGTTGGTGGTGGAC 1139

```

2217 GACCTGGACGGCTGCCCAACCTCAATCTGRTCTGCGCCACCAACGCCACCTACCACTGC 2276
 2125 GACCTGGACGGCTGCCCTAAATGAAAACCTGGTGTGTGTGGCCAAACCAACCTACCACTGC 2066
 2277 ATCAAGGATTAAGTCCCTATCTGCAATCTGGCCAGGAGACTTTGACAGAGCGGG 2336
 2065 AAAAAGGCAACTGCCCCCAACCTTCCCACTCGGGCAGAGACTATGACAAGGACGGG 2006
 2337 ATTGGCGATGCTGTGATGATGACGATGACAACTGAGCGGTGTGACCGATGAGAAAGGCAAC 2396
 2005 ATTGGCGATGCTGTGATGATGACGATGACAACTGAGCGGTGTGACCGATGAGAAAGGCAAC 1946
 2397 TGGCAGCTCTCTCAATCCCGCCGAGGCTGACTATGACAAGGATGAGTTGGGACCGC 2456
 1945 TGTCCATTTCAATTACCCAGCCAGTATGACTATGACAGAGATGATGTGGAGACCGC 1886
 2457 TGTGACAACTGCCCTTACGTGACCAACCTGCCAGATCGACAGACAACTATGAGAG 2516
 1885 TGTGACAACTGCCCTTACCAACCAACCTGACCAAGCAGACAGACAAACCGGGAG 1826
 2517 GGTGACGCTGCTCGTGGACATTTGATGGGACGATGCTTCAATGAAAGGACAAATTGT 2576
 1825 GGCATGCTGCTGTGTGACATCGATGAGATGGAATCCTCAATGAACGAGACAACTGC 1766
 2577 CCCTACGTCTACAACTGACAGGAGGACACCGATGGTCAACGCTGTGGGGATCACTGT 2636
 1765 CAGTACGTTTACAACTGACAGGAGGACACCGATGGGTGGAGATCAGTGT 1706
 2637 GACAACTGCCCTGCTGACAGCAACCTGACAGCCGATGACAGCGGATGACCTTGTGGG 2696
 1705 GACAACTGCCCTGCTGACAGCAACCTGACAGCCGATGACCTGACCTGACCTCATAGGG 1646
 2697 GACCACTGTGACAAACAGGAGGACATAGATGACGAGCGGACCAAGCAACAGGACCAAC 2756
 1645 GACAACTGTGACAAACATCAGGACATCGATGAGGATGGCCATCAGAACAACTCTGGACAC 1586
 2757 TGCCCTTACATCTCAACGCGCAACAGGCTGACCATGACAGAGAGCGGCGAGCGCC 2816
 1585 TGTCCCTATGTGCTTAATGCGCAACAGGCGGACCATGATTAAGATGGCAAGAGATGCC 1526
 2817 TGTGACCTGATGACAAACGATGCGCTGCCGATGACAGGAGCAACTGCGGCTGTG 2876
 1525 TGTGACCATGACGATGACAAAGCGCATCTCTGATGACAGAGCAACTGCGGCTGTG 1466
 2877 TTCAACCCAGACCGAGGAGCTTGGACGCTGATGACGCGGTGATTTGTAAAGATGAT 2936
 1465 CCCAATCTGACCAAGGAGCTTGTATGATGATGCGCGGATGACGCTGCAAGAGCGAC 1406
 2937 TTTGACATGACAACTCCAGATATTGATGATGATGCTGCTGAAACAACTGCAATCAGT 2996
 1405 TTTGACATGACAACTGCGCAGATATTGATGATGATGCTGCTGAGAAATTTGACATCAGT 1346
 2997 GAGACAGACTTCAGGAACTTCAGATGCTGCTGATGATGATGATGATGATGATGATGATGAT 3056
 1345 GAAACCGATTTCCGACAAATTCAGATGATTTCTAGATGATGATGATGATGATGATGATGAT 1286
 3057 GATCCCACTGGCTATTCGCACTCAAGCAAGGAGCTGTTTCCAGACAGCAACTCGGAC 3116
 1285 GACCTTAATCTGGTGTGCGCCATCAGGCGAAGAACTTGTCCAGACTGTAACTGTGAC 1226
 3117 CCGGCACTGCTGATGATTTTCCAGATTTGGGTCTGTGAGCTTCACTGAGTCACTTCTAC 3176
 1225 CCGGACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1166
 3177 GTAAACACTGACCGGAGCAGCACTATCGCGCTTCTGTTGGTTACCACTGCAAGCAGC 3236
 1165 ATCAACCCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1106
 3237 CGCTTCTATGTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3296
 1105 CGCTTCTACGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1046

QY 3297 CGGCGCTATGGCTACTCCGGCGTGTCTCTCAAGGTGGTGAACCTCCACCGGGACCGGC 3356
 Db 1045 AGGGCTCAGGGATACTCAGCGCTGTCTGTAAAGTTGTAACTCCACACCGGCGCTGCG 986
 QY 3357 GAGCACCTGAGAAACGCGCTGTGGCACACGGGAAACACCGCGGAGGTGCGAACCTTA 3416
 Db 985 GAGCACCTGCGGAATGCACTGTGGCACACAGGAAACACCGCTGCGGAGGTGCGAACCTG 926
 QY 3417 TGGCACGACCCAGGAAACATTTGGCTGGAAAGGACTACACGGCCTATAGTGGCACTGACT 3476
 Db 925 TGGCATGACCTCGCCACATCGCTGGAAGATTTTCACTGCATACAGATGCGCTCTCAGC 866
 QY 3477 CACAGGCCAAGACCTGCTACATCAGAGTCTTAGTGCATGAGGAAACAGGTTCATGSCA 3536
 Db 865 CACAGGCCAAGACCTGCTATATCAGAGTGGTGTATGAGGAAAGAAATCATGCTC 806
 QY 3537 GACTCAGGACCTATCTATGACCAACCTACGCTGGCGGCGGCTGGTCTTATTTGTCTTC 3596
 Db 805 GACTCGGAGCCATCTATGACAAACCTACGCGGCGGTAGACTAGGCTGTTCCTTC 746
 QY 3597 TCTCAGAAATGGTCTATTTCTCAGACCTCAAGTACGATGCAAGATTTTAAACAAGA 3656
 Db 745 TCTCAGAAATGGTCTTCTCTCAGACATGAATACGATGTCGAGATTCCTATCATCA 686
 QY 3657 TTTGC 3661
 Db 685 GCTGC 681

RESULT 6
 LOCUS AY405252 3412 bp DNA linear GSS 16-DEC-2003
 DEFINITION Homo sapiens THBS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCSSION AY405252
 VERSION AY405252.1 GI:39761226
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3412)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 3412)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submision
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment
 FEATURES
 Location/Qualifiers
 1..3412
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>3412
 /gene="THBS1"
 /locus_tag="HCM2162"

gene
 source
 1..3412
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>3412
 /gene="THBS1"
 /locus_tag="HCM2162"

ORIGIN
 Query Match 23.9%; Score 1360.8; DB 29; Length 3412;
 Best Local Similarity 61.2%; Pred. No. 1.2e-193;
 Matches 2142; Conservative 0; Mismatches 1243; Indels 114; Gaps 3;


```
1074 AACTCAGGCGAGGAGACTATGACAAAGATGGAATTGGTGATGCCCTGTGATGATGACGAT 2133
1075 GACAAATGACGGTGTGACCGATGAGAGGACAACTGCCAGCTCTCTTCAATCCCGCCAG 2423
1076 GACAAATGATAAATTTCCAGATGACAGGGGACAACTGTCCATTCATTACAAACCCAGCTCAG 2193
1077 GCTGACTATGACAGGATGAGGTTGGGACCGCTGTGACAACTGCCCTTACGTGACCAAC 2483
1078 TATGACTATGACAGAGATGATGTGGGAGCCGCTGTGACAACTGTGCCCTTACCAACCAAC 2253
1079 CCTGCCAGATTCGACACAGACAAATGAGAGGGGTGAGCCCTGTGCCGTGACACATTGAT 2543
1080 CCAGATCAGGCAGACACAGACAAATGGGAGGAGAGCCCTGTGCTCGACACATTGAT 2313
1081 GGGAGCATGCTTCAATGACAGGACAAATTTGCTCCCTAGCTGTACAACTGACACAGAGG 2603
1082 GGAACCGGTATCTCAATGAAACGGGACAACTGCCAGTAGCTTACAAATGTGACCAAGAG 2373
1083 GACACGGATGCTGACGGTGTGGGGATCACTGTGACAACTGCCCTGTGACCAACCTT 2663
1084 GACACTGATATGATGGGTTGGAGATCAGTGTGACAACTTGCCCTTGGAAACACATCCG 2433
1085 GACAGACCGAGCTGGACAAATGACCTTTGGGAGCCAGTGTGACAAACAGGAGCAATA 2723
1086 GATCAGCTGACTCTGACTCAGACCGCATTTGGAGATACCTGTGACAAACATCAGGATAT 2493
1087 GATGACACGGCCACAGAACAAACAGGACAACTGCCCTTACATCTCCAAACGCCAACAG 2783
1088 GATGAAGATGGCCACAGAACAACTGTGACAACTGTCCCTATGTGCCAATGCCAACAG 2553
1089 GCTGACCATGACAGAGAGCGGACGAGCGCTGTGACCTGTGACCTGTGATGACAAACGATGGC 2843
1090 GCTGACCATGACAAAGATGGCAAGGAGATGCTGTGACACGATGATGACAAACGATGGC 2613
1091 GTCCCGATGACAGGGACAACTGCCGCTGTGTTCAACCCAGACAGGAGACTTGGAC 2903
1092 ATTCCTGATGACAGGACAACTGCACTGTGCCAATCCGACGAGGACTCTGAC 2673
1093 GGTGATGACGGGTGATATTTGTAAGATGATTTGACAAATGACAAATCCAGATAT 2963
1094 GGCATGCTGAGGTGATGCTGCAAGATGATTTTACCATGACAGTGTGCCAGACATC 2733
1095 GATGATGTGTCTGAAACAACTGCCATCAGTGAGACAGATTCAGGAACTCCAGATG 3023
1096 GATGACATCTGCTGAGATGTTGACATCAGTGAGACCGATTTCCGCGATTCAGATG 2793
1097 GTCCCTTGGATCCCAAGGACCAACCAAAATGATCCCAACTGGGTCTATCGCCATCAA 3083
1098 ATTCCTCTGGACCCCAAGGACATCCCAAAATGACCTTAACCTGCTTGTAGCCATCAG 2853
1099 GGCAGGAGCTGGTTCAGACAGCAACTCGGACCCGGCATCGCTGTAGTTTGACGAG 3143
1100 GGTAAAGAACTCGTCCAGCTGTCAACTGTGATCTGGACTCGCTGTAGTTATGATGAG 2913
1101 TTGGGCTGTGGACTTCAGTGGCACTTCTACGTAAACACTGACCGGACGACGACTAT 3203
1102 TTTAATGCTGTGGACTTCAGTGGCACTTCTTCAATCAACACCGAAGGACGATGACTAT 2973
1103 CGCGCTTGTCTTTGTTTACAGTCAAGCAGCGCTTCTATGTGTGATGTGGAAGCAG 3263
1104 GCTGGATTTGTCTTTGGTACCAAGTCCAGTCCAGCAGCGCTTTTATGTTGTGATGGAAGCAA 3033
1105 GTGACCGAGACCTACTGGGAGGACAGCCAGCCAGCGGCTATGGCTACTCCGCGGTGTC 3323
1106 GTACCCAGCTCTACTGGGACACCAACCCACGAGGGCTCAGGGATCTCTGGGCCCTTCT 3093
1107 CTCAGGTGTGAACTCCACACGGGGAGCGGAGACCTGAGGAACCGCTGTGGCAC 3383
1108 GTGAAAGTTGTAACTCCACACAGGGCTGGGAGCACTCGGAAACCGCTGTGGCAC 3153
1109 ACGGGGAACACCGGGGACAGTGGGAACTTTATGACAGACCCCAAGGAACATTTGCTGG 3443
```

```
3154 ACAGGAACACCCCTGGCCAGGTGGCCACCTGTGTGATGACCCCTCGTCACATAGGCTGG 3213
3444 AAGGACTACACGGGCTATAGGTGGCACCTGACTCAGAGCCCAAGACTGGCTACATCAGA 3503
3214 AAGATTTCACGGCTACAGATGGGCTCTCAGCCACAGCCAAAGACGGGTTTCATTAGA 3273
3504 GTCTTAGTCATGAGGAAACAGGTCATGGCAGACTCAGGACCTATCTATGACCAAC 3563
3274 GTGGTGTATGATGAGGAGAAATCATGGTGTGACTCAGGACCCATCTATGATAAAAC 3333
3564 TACCTCGGCGGCGCTGGGCTATTTGTCTTCTCTCAAGAAATGGTCTATTTCTCAGAC 3623
3334 TATGCTGGTGTAGACTAGGTTGTTGTCTTCTCTCAAGAAATGGTGTCTTCTCTGAC 3393
3624 CTCAGTACGAATGCAGG 3642
3394 CTGAATACGAATGTAGAG 3412

RESULT 7
AY405254 3240 bp DNA linear GSS 16-DEC-2003
LOCUS
DEFINITION Mus musculus THBS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
ACCESSION AY405254
VERSION AY405254.1 GI:39761228
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3240)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3240)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..3240
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>3240
/gene="THBS1"
/locus_tag="HGM2162"

Query Match 21.4%; Score 1219; DB 29; Length 3240;
Best Local Similarity 61.4%; Pred. No. 1.9e-172;
Matches 2071; Conservative 0; Mismatches 1086; Indels 216; Gaps 3;

QY 270 GGGCCGACCCGGGGTCCGGCTTACCGCTTCGTGGCTTTTCTACTACATCCCAACCGGTG 329
DB 84 GGCCAGATCTATCCAGCCCGCTTCCGATGAGATGCCAACCCTGATCCCGCTGTG 143
QY 330 AACCGATGACCTCAGCAAGATCACAGATCATCGGAGAGGAGGCTTCTTCTC 389
DB 144 CCGGATGACAAGTTCGAAGACCTACTGGACGCTGTGTGGCGCCGACAAAGGCTTCATCTTC 203
QY 390 ACGCCCCAGCTCAGCAGGACGGCAAGTCCAGGGGACGCTGTGTGGCTCTGAGGGGCCCC 449
```



```
2610 GATGGTGACGGTGTGGGGATCACTGTGACACTGCGCCCTGCTGTCACACACCTGACCAG 2669
2620 GACATGATGGGTGGAGATCAGTGTGACAACTGCCCCCTGGAACACAACTCCAGACCAG 2667
2670 ACCAGCGTGGCAATAGACTTGTGGGACCAAGTGTGACAAACAGGAGCATPAGATGAC 2729
2680 CTGACTCTGACTCAGACCTCATAGGGGACACTTGTGACAACTCAGGACATCGATGAG 2327
2730 GACGGCCACCAAGAACACAGGACAACTGCCCTACATCTCCAAACCCACACAGGTGAC 2789
2820 GATGGCCATCAGACAACTGTGGACAACTGCTCCCTATGCTTAAACCCACAGGCGGAC 2387
2790 CATGACAGAGACGGCCAGGCGGACGCTGTGACCACTGATGATGACAAAGATGCGTCCCC 2849
2880 CATGATAAGATGCAAGAGGATGCTGTGACCACTGATGACCAATGACGCGCATCCCT 2447
2850 GATCACAGGAGCAACTGCGCGCTGTGTTCAACCCAGACAGGAGACTTGGACGGTGAT 2909
2448 GATCACAGAGACAACTGCGGCTGTGCTCCCAATCCTGACCAAGGACTCTGATGTGAT 2507
2910 GGACGGGTGATATTGTAAGATGATTTGACATGATGACAACTCCAGATATTGATGAT 2969
2508 GCGCGAGGTGACGGCTGCAAGAGACACTTTGACCAATGACAATGTGCCAGATATTGATGAC 2567
2970 GTGTGTCTCTGAAAAACAATGCCATCAGTGAGACAGACTTCAGGAACCTTCAGATGTTGGG 3029
2568 ATCTGTCTCTGAGATTTTGACATCAGTGAACCGATTTCGACGATTCAGATGATTCCT 2627
3030 TTGGATCCCAAGGACCAACCAATTCATCCAACTGGGTATTCGCGCATCAGGCAAG 3089
2628 CTGATCCCAAGGACCACTCCCAATGACCTTAACCTGCGGTGTTGCGCCATCAGGGCAA 2687
3090 GAGCTGTTTCAGACAGCAGCAACTCGGACCCGCGCATCGCTGTAGTTTGTGACAGTTTGGG 3149
2688 GAATCTGCTCAGACTGTAACTGTGACCTGACTGCTGTAGTTGATGATGATTTAAT 2747
3150 TCTGTGACTTCAGTGGGACATTTCTACGTAACTGACCGGACGACCACTATGCGGC 3209
2748 GCTGTGACTTCAGCGGTACCTTTCTATCAACCGAGAGAGATGATGACTACGCTGGC 2807
3210 TCGTCTTTGTTTCCAGTCAAGCAGCGCTTCTATGTGTGATGTGGAGCAGGTGACG 3269
2808 TTTGTTATTCGGCTACCACTGACGAGCGCTTCTACGTTGTGATGTGGAACAAGTCACC 2867
3270 CAGACCTACTGGGAGGACCGCCACGCGGCTATGCTACTCGGCGTGTCCCTCAAG 3329
2868 CAGTCTCTAGGGACCAACCCCAAGGGCTCAGGGATATCTAGGCGCTGTCTGTAAG 2927
3330 GTGTGAATCCACACGGGACCGGCGAGCACTGAGGAACCGCTGTGGCACACGGGG 3389
2928 GTTGTGAATCCACACCGGCGCTGGCGAGCACTGCGGAATGCACTGTGGCACACAGGA 2987
3390 AACACGCGGGGAGGTGCGAACCTTATGGCACGACCCCGAGGAACATTCGGCTGGAGGAC 3449
2988 AACACCTCTGGCCAGGTGCGACCTGTGGGATGACCTTGGCCACATCGCTTGGAAAGAT 3047
3450 TACACGCGCTTATAGGTGCACTGACTCAGCGGCCAAGACTGGCTTACATCAGAGCTTA 3509
3048 TTTCACTCGTACAGATGCGCTCTAGCCACAGCGCCAAAGACCGGTTATATCAGAGTGGT 3107
3510 GTGCATCAAGGAACAAGCTCATGTCAGACTCAGGACCTTATGACCAACACCTACGCT 3569
3108 ATGTATGAAGGAAGAAATCATGTGCTACTCGGACCCATCTATGACAAACCTACGCC 3167
3570 GCGGGCGGTGCTATTTGCTCTCTCAAGAAATGGTCTATTTCTCAGACCTCAAG 3629
3168 GCGGTGAGACTAGCGCTGTCTCTCTCAGGAAATGGTGTCTCTCTCAGACATGAA 3227
3630 TACGAATGACAG 3642
3228 TACGAGTGTGAG 3240
```

```
RESULT 8
BQ933301
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
Source
1..939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6205936"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski sciatic nerve"
/notes="Vector: pCMV-SF0R16 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACCGCTCCG-3' and
5'-GACGTAGTCTAGATCGGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
```

```
BQ933301 939 bp mRNA linear EST 21-AUG-2002
AGENCOURT 8744271 Lupski_sciatic_nerve Homo sapiens cDNA clone
IMAGE:6205936 5', mRNA sequence.
BQ933301
BQ933301.1 GI:22348684
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13629 row: j column: 17
High quality sequence stop: 686.
Location/Qualifiers
1..939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6205936"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski sciatic nerve"
/notes="Vector: pCMV-SF0R16 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACCGCTCCG-3' and
5'-GACGTAGTCTAGATCGGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
```

ORIGIN

```
Query Match 15.0%; Score 854; DB 13; Length 939;
Best Local Similarity 97.6%; Pred. No. 9.1e-118;
Matches 919; Conservative 0; Mismatches 17; Indels 6; Gaps 5;

QY 4706 CTAGTAACACATATGTAGTATAATTTCTAGAACAAACATCTTAATAGTATATAATCTG 4765
Db 1 CTAGTAACACATATGTAGTATAATTTCTAGAACAAACATCTTAATAGTATATAATCTG 60
QY 4766 TGAATAATATAGGCTTGATAATATTAGGTGTGTCAGCATGAAGCATGCTAGAAGCTGTAAC 4825
Db 61 TGAATAATATAGGCTTGATAATATTAGGTGTGTCAGCATGAAGCATGCTAGAAGCTGTAAC 120
QY 4826 AGATACATAGAGATATAGGAGTTTATGATGGAACCTTAATATATATATATCTGTCAG 4885
Db 121 AGATACATAGAGATATAGGAGTTTATGATGGAACCTTAAATATATATATATCTGTCAG 179
QY 4886 CGATTTAGTTCAATATTGTTACTGTTATCTATCTGCTGTTATATGGAATCTTTTAATT 4945
Db 180 CGATTTAGTTCAATATTGTTACTGTTATCTATCTGCTGTTATATGGAATCTTTTAATT 239
QY 4946 CAACGCTGAAA-GATCAGCATTTAGTCTGCCAGGCACACCCCAATATCATCATGTATGT 5004
```



```

Db 541 GTATCTAGCCTTTGGTAACCTGGGACAAAGTTGCTTTTACCTGATTTGATGATATCA 600
Qy 4617 TTAAGTTCCAGTTATAAATATTTTCTTAATATTTAATTAAGTGAATGCAACTC 4676
Db 601 TTAAGTTCCAGTTATAAATATTTTCTTAATATTTAATTAAGTGAATGCAACTC 660
Qy 4677 CATTTACAGTAACCTTATTTTAAATATGCTAGTAAACACATATGTAATTTCTAGA 4736
Db 661 CATTTACAGTAACCTTATTTTAAATATGCTAGTAAACACATATGTAATTTCTAGA 720
Qy 4737 AACAACTCTAATTAAGTATATAATCCTGTGAAATATAGG 4778
Db 721 AACAACTCTAATTAAGTATATAATCCTGTGAAATATAGG 762

RESULT 14
BQ940133 904 bp mRNA linear EST 21-AUG-2002
AGENCOURT 8784556 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6376342
5', mRNA sequence.
BQ940133
BQ940133.1 GI:22355611
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 904)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: AFCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC2558 row: f column: 23
High quality sequence stop: 551.
Location/Qualifiers
1. 904
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6376342"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_43"
/notes="Organ: eye; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

Query Match 12.9%; Score 736; DB 13; Length 904;
Best Local Similarity 99.6%; Pred. No. 3.9e-100;
Matches 747; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 2342 CGATGCTGTGATGACGATGACATGACGGTGTGCCATGAGAGGACAACTGCCA 2401
Db 1 CGATGCTGTGATGACGATGACATGACGGTGTGCCATGAGAGGACAACTGCCA 60
Qy 2402 GCTCCTCTTCAATCCCGCCAGGCTGACTATGACAAGGATGAGGTTGGGACCGCTGTGA 2461
Db 61 GCTCCTCTTCAATCCCGCCAGGCTGACTATGACAAGGATGAGGTTGGGACCGCTGTGA 120

```

```

Qy 2462 CAACTGCCCTTACGTGCACAAACCTCTGCCAGATCGACACAGACAACAACTGAGAGGTGA 2521
Db 121 CAACTGCCCTTACGTGCACAAACCTCTGCCAGATCGACACAGACAACAACTGAGAGGTGA 180
Qy 2522 CGCCTGCTCCGTGGACATTTGATGGGACGATGCTTCAATGAACGAGACAATTTGTCCTTA 2581
Db 181 CGCCTGCTCCGTGGACATTTGATGGGACGATGCTTCAATGAACGAGACAATTTGTCCTTA 240
Qy 2582 CGCTTACAACTGACACGAGGACACGGATGTCAGGTGTGGGGATCACTGTGACAA 2641
Db 241 CGCTTACAACTGACACGAGGACACGGATGTCAGGTGTGGGGATCACTGTGACAA 300
Qy 2642 CTGCCCCCTGTGTGCACAAACCTTGACAGACCGAGTGAACAATCTTTGTTGGGACCA 2701
Db 301 CTGCCCCCTGTGTGCACAAACCTTGACAGACCGAGTGAACAATCTTTGTTGGGACCA 360
Qy 2702 GTGTGACACAAACGAGGACATAGATGACGACGGCCACAGAAACAACAGGACAACCTGCC 2761
Db 361 GTGTGACACAAACGAGGACATAGATGACGACGGCCACAGAAACAACAGGACAACCTGCC 420
Qy 2762 CTACATCTCCAAACGCAACCGAGCTGACCATGACAGAGACGGCGGCGAGCGCTGTGA 2821
Db 421 CTACATCTCCAAACGCAACCGAGCTGACCATGACAGAGACGGCGGCGAGCGCTGTGA 480
Qy 2822 CCTGTATGATGACAAACGATGGCTGCCCGATGACAGGACAACTGCCGGCTTGTGTTCAA 2881
Db 481 CCTGTATGATGACAAACGATGGCTGCCCGATGACAGGACAACTGCCGGCTTGTGTTCAA 540
Qy 2882 CCAGACCAAGGAGGACTTGGACGGTGTGATGACGGGGTGATATTGTTAAAGATGATTTGA 2941
Db 541 CCAGACCAAGGAGGACTTGGACGGTGTGATGACGGGGTGATATTGTTAAAGATGATTTGA 600
Qy 2942 CAATGACACATCCCAGATATTGATGATGTGTCTTGAAACAAATGCGATCAGTGAGAC 3001
Db 601 CAATGACACATCCCAGATATTGATGATGTGTCTTGAAACAAATGCGATCAGTGAGAC 660
Qy 3002 AGACTTCAGAACTTCCAGATGTGCCCTTGGATGCCAAAGGGACCA-CCCAAATGATC 3060
Db 661 AGACTTCAGAACTTCCAGATGTGCCCTTGGATGCCAAAGGGACCA-CCCAAATGATC 720
Qy 3061 CCAACTGGGTGCTTCCGCTCATCAAGGCAAGG 3090
Db 721 CCAACTGGGTGCTTCCGCTCATCAAGGCAAGG 750

RESULT 15
BQ618192/c
LOCUS BQ618192
DEFINITION UI-H-FHI-bfe-k-04-0-UI.s1 NCI CGAP FHI Homo sapiens cDNA clone
UI-H-FHI-bfe-k-04-0-UI.3', mRNA sequence.
BQ618192
BQ618192.1 GI:23284407
ACCESSION BQ618192
VERSION Homo sapiens (human)
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 750)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 637-685, >(CAAA)n#Simple_repeat

```


Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES
source

Location/Qualifiers
1..750

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-H-FH1-bfe-k-04-0-UI"

/tissue_type="Cell Line"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP_FH1"

/note="Organ: Chondrosarcoma; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_FH1 is a normalized cDNA library obtained from a cell line derived from grade I chondrosarcoma tissue. The library was constructed and normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGAATCCGGC. The cell line was provided by Dr. James Martin from the University of Iowa."

TAG_LIB=UI-H-FH1
TAG_SEQ=AGAATCCGGC"

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

Chondrosarcoma

RIGIN

Query Match

Best Local Similarity 12.9%; Score 735.6; DB 13; Length 750;

Matches 738; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Y 4955 AAAAGAAATCAGCAATTTAGCTTGGCAGGACACACCAATAATCAGTCATGTGTAATAGCA 5014

b 748 AACAGAAATCAGCAATTTAGCTTGGCAGGACACACCAATAATCAGTCATGTGTAATAGCA 589

Y 5015 CAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 5074

b 688 CAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 629

Y 5075 ATGATCTTCTGCGAGGAATAGTCACTCATCCACTCCACATAAGGGGTTTAGTAAGAGA 5134

b 628 ATGATCTTCTGCGAGGAATAGTCACTCATCCACTCCACATAAGGGGTTTAGTAAGAGA 569

Y 5135 AGTCGTCTGTCGATGATGATAGGGGGCAATCTTTTCCCTTCTGTTAATAGTCA 5194

b 568 AGTCGTCTGTCGATGATGATAGGGGGCAATCTTTTCCCTTCTGTTAATAGTCA 509

Y 5195 TCACATTTCTATGCGAAACAGGAAACATCCATPACTTAGTCTTAATGATACATTCAT 5254

b 508 TCACATTTCTATGCGAAACAGGAAACATCCATPACTTAGTCTTAATGATACATTCAT 449

Y 5255 TTTGATATAAATTTTGTGTTTCCCTTGGAGTTGATCGTTGTTGTTGTTGTTGTTGTTGTTG 5314

b 448 TTTGATATAAATTTTGTGTTTCCCTTGGAGTTGATCGTTGTTGTTGTTGTTGTTGTTGTTG 389

Y 5315 CACTTTTACTTTTTTGGGTGGAGCTGTATCCCGAGACCAACGAGCGTTGGGATAC 5374

b 388 CACTTTTACTTTTTTGGGTGGAGCTGTATCCCGAGACCAACGAGCGTTGGGATAC 329

Y 5375 TTCAATTAATAGCGAGTGTCAACAGCGGTGAGGTTTTCTGTTCTGTTGTTGTTGTTGTTGTTG 5434

b 328 TTCAATTAATAGCGAGTGTCAACAGCGGTGAGGTTTTCTGTTCTGTTGTTGTTGTTGTTGTTG 269

Y 5435 AACCGTACAATGGTGTGGAGTGACGATGATGTAATTTAGAAATGTAACCATATTTTTT 5494

b 268 AACCGTACAATGGTGTGGAGTGACGATGATGTAATTTAGAAATGTAACCATATTTTTT 209

QY 5495 GTAAATTTATGTTTCTTAAACABATTTATCGTATAGGTTGATGAAACGTCATGTGT 5554
Db 208 GTAAATTTATGTTTCTTAAACABATTTATCGTATAGGTTGATGAAACGTCATGTGT 149
QY 5555 TTTGCCAAAGACTGTAAATATTTTATTTATGTTTCAATGGTCAAAAATTTTCAACACTGAA 5614
Db 148 TTTGCCAAAGACTGTAAATATTTTATTTATGTTTCAATGGTCAAAAATTTTCAACACTGAA 89
QY 5615 ACCCTGCACCTTAGCTAGAACCTCATTTTTTAAAGATTAAACAACAGGAATTAATTTGTAATA 5674
Db 88 ACCCTGCACCTTAGCTAGAACCTCATTTTTTAAAGATTAAACAACAGGAATTAATTTGTAATA 29
QY 5675 AAGCTTTTCTAAAAAATAAAAAA 5696
Db 28 AAGCTTTTCTAAAAAATAAAAAA 7

Search completed: March 11, 2004, 14:23:35

Job time : 9447.57 secs

Klausner, R.D., Collins, F.S., Wagner, L., Shennen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Moore, H., Max, S.I., Wang, J., Hsieh, F., Hopkins, R.F., Jordan, H., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Diatchenko, L., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Stapleton, M., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Scheetz, T.E., Frange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Carninci, P., Mullah, S.J., Bosak, S.A., McEwan, P.J., Abramson, R.D., Malek, J.A., Gunaratne, P.H., Richards, S., McKernan, K.J., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Worley, D.C., Kuzny, D.M., Sodergren, E.J., Ku, X., Gibbs, R.A., Vahney, J., Helton, E., Kettenan, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Buttefield, S., Krzywinski, M.I., Skalska, U., Smalus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2388257

1247732

2 (bases 1 to 1938)

Strausberg, R.

Direct Submission

Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: amadan@systemsbiology.org

Anup Madan, Jessica Fahey, Erin Halton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Place: 19 Row: k Column: 23

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505594.

Location/Qualifiers

1..1938

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:13616 IMAGE:4281085"

/tissue_type="Brain, Primitive neuroectodermal"

/clone_lib="NIH MGC_58"

/lab_host="DH10B"

/note="Vector: pDNR-LIB"

1..1938

/gene="SERPINE2"

/note="synonyms: PAI2, PLANH2, PAI"

/db_xref="LocusID:5055"

/db_xref="MIM:173390"

80..1327

/codon_start=1

/product="serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 2"

/protein_id="AAH12609.1"

/db_xref="GI:15214956"

/db_xref="LocusID:5055"

/translation="MEDLCVANTLFAALNFKHLAKSPQNLFLSPWSISSTWAVYVSGRSTEDQMAKVLQNEVGANAATVPENFTSCGFMQIQKGSYPDAILOQAADKIHSSFRSLDANSTGNLLESVKNLFKGSASFREYIRLCQKYSSEPPQVDFLECAEARKKIYSVKTQTKGINLLPEGSVDGTRVLI VNAVYFKGWKTPPEFKLNGLYPFRVNSAQRTPVQMWYLRKLNIGVIEDLKAQILLELPVAGDVSFELLPLDELADVS

TGLELESELTVDKLNKWTSKDQMAEDVEVYIPOFKLEHVELRILSRMGMDAPN
XGRANFGMSERNDLFLSEVPHQAVDNEEGTEAAAGTGGVMTGTHGGPQFVADH
PFLFLNHKTIINCILFFGRFSP"

116..1324

/note="SERPIN; Region: SERine Proteinase Inhibitors"

/db_xref="CDD:smart00093"

ORIGIN

Query Match 99.3%; Score 1894.8; DB 9; Length 1938;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GAGGAGCATTCGCCGTGACAGACGACCACTCAGAGAATAACACAGAGAACCAACCATTCGAAA 60
Db 18 GAGGAGCATTCGCCGTGACAGACGACCACTCAGAGAATAACACAGAGAACCAACCATTCGAAA 77
Qy 61 CAATGGAGGATCTTTGTGTGGCAAAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGG 120
Db 78 CAATGGAGGATCTTTGTGTGGCAAAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGG 137
Qy 121 CAAAAGCAAGCCCAACCCAGAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGG 180
Db 138 CAAAAGCAAGCCCAACCCAGAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGG 197
Qy 181 CCATGGTCTATATGGGTCGAGGCGACGACGAGACGACGATGCGCAAGTCTTCAGT 240
Db 198 CCATGGTCTATATGGGTCGAGGCGACGACGAGACGATGCGCAAGTCTTCAGT 257
Qy 241 TTAATGAAGTGGGAGCCAAATGAGTTACCCCATGATGATGCGGATGCGGACGACGATG 300
Db 258 TTAATGAAGTGGGAGCCAAATGAGTTACCCCATGATGATGCGGATGCGGACGACGATG 317
Qy 301 GGTTTCATGACGAGATCCAGAGGGGTAGTTATCTTCATGCGATTTTCAGGCGACAAAGCTG 360
Db 318 GGTTTCATGACGAGATCCAGAGGGGTAGTTATCTTCATGCGATTTTCAGGCGACAAAGCTG 377
Qy 361 CAGATAAATTCATTCCTTCGCTCTCAGCTCTGCAATCAATGATGATGATGATGATGATG 420
Db 378 CAGATAAATTCATTCCTTCGCTCTCAGCTCTGCAATCAATGATGATGATGATGATGATG 437
Qy 421 ATTATTTACTGGAAAGTGTCAATAAGCTGTTTGGTGAGAAAGTCTGCGAGTTCGCGGAAG 480
Db 438 ATTATTTACTGGAAAGTGTCAATAAGCTGTTTGGTGAGAAAGTCTGCGAGTTCGCGGAAG 497
Qy 481 AATATTTTCGACTCTGTCAGAAATATTACTCTCAGAACCCAGGAGTAGATCTTCCTAG 540
Db 498 AATATTTTCGACTCTGTCAGAAATATTACTCTCAGAACCCAGGAGTAGATCTTCCTAG 557
Qy 541 AATGTCAGAGAGAGCTAGAAAAAGATTAATTTCTCGGTCAAGACTCAAAACCAAGGCA 600
Db 558 AATGTCAGAGAGAGCTAGAAAAAGATTAATTTCTCGGTCAAGACTCAAAACCAAGGCA 617
Qy 601 AAATCCCAAACTTGTACTCTGAAGGTTCTGTAGATGGGATACAGGATGGTCTCGTGA 660
Db 618 AAATCCCAAACTTGTACTCTGAAGGTTCTGTAGATGGGATACAGGATGGTCTCGTGA 677
Qy 661 ATGCTGTCTACTTCAAGAGAAAGTGGAAAACTCCATTGTGAGAGAACTAAATGGGCTTT 720
Db 678 ATGCTGTCTACTTCAAGAGAAAGTGGAAAACTCCATTGTGAGAGAACTAAATGGGCTTT 737
Qy 721 ATCTTTCCGTGTAAACTCGGCTCAGCGCACACTCTGACAGATGATGATCTTGTGCGTGA 780
Db 738 ATCTTTCCGTGTAAACTCGGCTCAGCGCACACTCTGACAGATGATGATCTTGTGCGTGA 797
Qy 781 AGCTAAACATTCGATACATAGAGACCTAAAGGCTCAGATTTCTAGAACTCCCATATGCTG 840
Db 798 AGCTAAACATTCGATACATAGAGACCTAAAGGCTCAGATTTCTAGAACTCCCATATGCTG 857
Qy 841 GAGATGTTAGCATGTTCTTTGTTGCTCCAGATGAATTCGCGATGTTCCACTGGCTGG 900
Db 858 GAGATGTTAGCATGTTCTTTGTTGCTCCAGATGAATTCGCGATGTTCCACTGGCTGG 917
Qy 901 AGCTCTGGAAAGTGAATAACCTATGACAAACTCAACAAAGTGGACCAAGCAAGACAAA 960

108064.1 GI:589223
Unknown.
Unclassified.
1 (bases 1 to 1920)
Webb,A.C. and Auron,P.E.
An arg-serpin human plasminogen activator inhibitor designated
PAI-2
Patent: EP 0278696-A2 1 17-AUG-1988;
Location/Qualifiers
1-1920
/organism="unknown"
/mol_type="unassigned DNA"

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Query Match 98.8%; Score 1886; DB 6; Length 1920;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1900; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 1 GAGGAGCATTGCCCGTCAGACAGCAACTCAGAGAAATACCCAGAGACACACCATTTGAAA 60
DB 11 GAGGAGCATTGCCCGTCAGACAGCAACTCAGAGAAATACCCAGAGACACACCATTTGAAA 70
QY 61 CAATGGAGGATCTTTGTGTGCAAAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGG 120
DB 71 CAATGGAGGATCTTTGTGTGCAAAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGG 130
QY 121 CAAAGCAAGCCCAACCCAGAGCACTCTTCTCTCCCATGGAGCATCTCGTCCACCATGG 180
DB 131 CAAAGCAAGCCCAACCCAGAGCACTCTTCTCTCCCATGGAGCATCTCGTCCACCATGG 190
QY 181 CCATGGTCTACATGGGCTCCAGGGGACGACCGAAAGACAGATGGCCAAAGTCTTCAGT 240
DB 191 CCATGGTCTACATGGGCTCCAGGGGACGACCGAAAGACAGATGGCCAAAGTCTTCAGT 250
QY 241 TTAATGAAGTGGGAGCCAAATGCAGTTACCCCATCACTCCAGAGAACTTTTACCAGCTGTG 300
DB 251 TTAATGAAGTGGGAGCCAAATGCAGTTACCCCATCACTCCAGAGAACTTTTACCAGCTGTG 310
QY 301 GGTTTCATGCGAGCAGATCCAGAAAGGTAGTTATCTGTAGTCGATTTTTCGAGGCAAGCTG 360
DB 311 GGTTTCATGCGAGCAGATCCAGAAAGGTAGTTATCTGTAGTCGATTTTTCGAGGCAAGCTG 370
QY 361 CAGATAAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGA 420
DB 371 CAGATAAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGA 430
QY 421 ATTATTTTCTGGAAGGTGTCAATAGCTGTTTGGTGAAGTCTGCGAGCTTCGGGAAG 480
DB 431 ATTATTTTCTGGAAGGTGTCAATAGCTGTTTGGTGAAGTCTGCGAGCTTCGGGAAG 490
QY 481 AATATATTTCCGACTCTGTCCAGAAATATTACTCTCAGAACCCAGGCGAGTAGACTTCCTAG 540
DB 491 AATATATTTCCGACTCTGTCCAGAAATATTACTCTCAGAACCCAGGCGAGTAGACTTCCTAG 550
QY 541 AATGTGCAAGAGAGCTAGAAAAGATTAATCTTGGTCAAGACTCAACCAAGGCA 600
DB 551 AATGTGCAAGAGAGCTAGAAAAGATTAATCTTGGTCAAGACTCAACCAAGGCA 610
QY 601 AAATCCCAAACTTGTACTCTGAAGGTTCTGTAGATGGGATACCCAGGATGGTCTCTGTGA 660
DB 611 AAATCCCAAACTTGTACTCTGAAGGTTCTGTAGATGGGATACCCAGGATGGTCTCTGTGA 670
QY 661 ATGCTGTCTACTTCCAAAGGAAGTGGAAAATCCCAATTTGAGAGAACTAAATGGGCTTT 720
DB 671 ATGCTGTCTACTTCCAAAGGAAGTGGAAAATCCCAATTTGAGAGAACTAAATGGGCTTT 730
QY 721 ATCTTTTCCGCTGTAACCTCGGCTCAGCGCACACCTGTACAGATGATGTACTTCGCTGAAA 780
DB 731 ATCTTTTCCGCTGTAACCTCGGCTCAGCGCACACCTGTACAGATGATGTACTTCGCTGAAA 790
QY 781 AGCTAAACATTTGGATACATAGAACCTTAAGGCTCAGATTCTTAGAACTCCCATATGCTG 840

RESULT 2
108064
LOCUS
DEFINITION
ACCESSION
Sequence 1 from Patent EP 0278696.
108064
PAT 02-DEC-1994
linear
1920 bp
DNA

Db 791 AGCTAAACATTTGATACATAGAGACCTAAAGGCTCAGATTCAGACTCCCATGCTG 850
|||
841 GAGATGTAGCATGTTCTTGTGCTTCCAGATGAAATTCGCGATGTGTCCACTGGCTTGG 900
|||
851 GAGATGTAGCATGTTCTTGTGCTTCCAGATGAAATTCGCGATGTGTCCACTGGCTTGG 910
|||
901 AGCTCTCGAAAGTGAATAAATCTATGACAAATCTCAACAGTGGACCAAGACAAAA 960
|||
911 AGCTCTCGAAAGTGAATAAATCTATGACAAATCTCAACAGTGGACCAAGACAAAA 970
|||
961 TGGCTGAGATCAAGTGTAGGTATACATACCCAGTTCAAATTAGAGAGACATTTATGAAC 1020
|||
971 TGGCTGAGATCAAGTGTAGGTATACATACCCAGTTCAAATTAGAGAGACATTTATGAAC 1030
|||
1021 TCAGATCCATTCGAAAGCATGGGATGAGAGAGCGCTTCAACAGGAGCGGCCCAATT 1080
|||
1031 TCAGATCCATTCGAAAGCATGGGATGAGAGAGCGCTTCAACAGGAGCGGCCCAATT 1090
|||
1081 TCTCAGGAGATGTCGGAGAGGAATGACTGTTCTTCTCTGAAGTGTTCACCAAGCCCATGG 1140
|||
1091 TCTCAGGAGATGTCGGAGAGGAATGACTGTTCTTCTCTGAAGTGTTCACCAAGCCCATGG 1150
|||
1141 TGGATGTGAATGAGGAGGCGCACTGAGACAGCGCTGCGACAGAGGTGTATGACAGGGA 1200
|||
1151 TGGATGTGAATGAGGAGGCGCACTGAGACAGCGCTGCGACAGAGGTGTATGACAGGGA 1210
|||
1201 GAACTGGACATGGAGGCCCAACAGTGTGCGCAGATCATCCGTTCTTTCTTATTATGC 1260
|||
1211 GAACTGGACATGGAGGCCCAACAGTGTGCGCAGATCATCCGTTCTTTCTTATTATGC 1270
|||
1261 ATAAGATAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCCTTAAACCTAAGCGT 1320
|||
1271 ATAAGATAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCCTTAAACCTAAGCGT 1330
|||
1321 GCTGCTTCGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAAT 1380
|||
1331 GCTGCTTCGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAAT 1390
|||
1381 TGCCAAAAATTTAGAGATGTTTTCTACATATTTCTGCTTTCTGAAACAATTTCTGTACC 1440
|||
1391 TGCCAAAAATTTAGAGATGTTTTCTACATATTTCTGCTTTCTGAAACAATTTCTGTACC 1450
|||
1441 CACTAAATTAACAACACAGAAATTAATAGCAATTTGCTATTAATTAACAATGACCAACCTATT 1500
|||
1451 CACTAAATTAACAACACAGAAATTAATAGCAATTTGCTATTAATTAACAATGACCAACCTATT 1510
|||
1501 AATCATTTGGTCTTCTAAATGGGATCATGCCATTTAGATTTTCTTACTATCAGTTTA 1560
|||
1511 AATCATTTGGTCTTCTAAATGGGATCATGCCATTTAGATTTTCTTACTATCAGTTTA 1570
|||
1561 TTTTATAACATTAACCTTTTACTTTGTTATTTATTTATTTATATAATGGTGAGTTTAA 1620
|||
1571 TTTTATAACATTAACCTTTTACTTTGTTATTTATTTATTTATATAATGGTGAGTTTAA 1630
|||
1621 ATATTGCTCAGTCCCTATTTAATGAGTCAATTAAGTTATAGAGCAGATGATCTGTTA 1680
|||
1631 ATATTGCTCAGTCCCTATTTAATGAGTCAATTAAGTTATAGAGCAGATGATCTGTTA 1690
|||
1681 ATTTCTCTATTAATAATGCTTTAATTTGTTCTCATTAATGAAGAATAAGTAGGTATCCCT 1740
|||
1691 ATTTCTCTATTAATAATGCTTTAATTTGTTCTCATTAATGAAGAATAAGTAGGTATCCCT 1749
|||
1741 CCATGCCCTCTGTAATAATCTGGAATAACATTAACATTAAGCAATATATATGTTA 1800
|||
1750 CCATGCCCTCTGTAATAATCTGGAATAACATTAACATTAAGCAATATATATGTTA 1809
|||
1801 TGTGCATTTCTAGAAATPACATAACATATATATGCTGTATCTTATTTCAATTGCAAG 1860
|||
1810 TGTGCATTTCTAGAAATPACATAACATATATATGCTGTATCTTATTTCAATTGCAAG 1869
|||
1861 TATATAATAATAAATCTGCTTCCAAACACAAAAATAAAAAA 1906
|||

Db 1870 TATATAATAATAAACCCTGCTTCCAAACACAAAAA 1915
|||
RESULT 3
A10352
LOCUS
DEFINITION Artificial gene fore plasminogenactivator-inhibitor type 2 (PAI-2).
ACCESSION A10352
VERSION A10352.1 GI:412084
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences
REFERENCE
1 (bases 1 to 2424)
AUTHORS
JOURNAL
FEATURES
location/Qualifiers
Patent: DE 3722673-A 1 19-JAN-1989;
1..2424
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
553..1800
/codon_start=1
/transl_table=11
/product="plasminogenactivator-inhibitor type 2 (PAI-2)"
/protein_id="CAA00891.1"
/db_xref="GI:412084"
/translation="MEDLCVANTFLALNLFPHLAKASPTQNLFLSPWSISSTMAVYM
GSRGTEQMAKVLQNEVGANAVTPTPENFTSCGFMQIQKSYDPAILQAADK
IHSPFSLSSAINASTGNLLEBSNKLFGESKASFREYIRLCOKYSESEPAVDLE
CAEARKINSWKTQTKGIPNLPEGSVDGTRMVLNVAVYFKGKWKTFEKLING
LYPPFVNSAORTPVOMMYLREKINIGVIEDLKAQILELPVAGDVSMFLLPDEIADVS
TGLELESISITDYKLNKWTSKDMADEVEVYIPOPKEBHYEYLRSLRSMGMEAFN
KGRANFSGMSERNDLPLSEVFHQAQVMDVNEEGTEAAAGTSGVMTGRTGHGPGPQFVADH
PFLFLMKHTNCTILFPGRFSP"
ORIGIN
Query Match 98.5%; Score 1880; DB 6; Length 2424;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1883; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 19 GACAGCAACTCAGAGAAATAACAGAGAACCAACAGATTGAAACAATGGAGGATCTTTGTG 78
DB 509 GACAGCAACTCAGAGAAATAACAGAGAACCAACAGATTGAAACAATGGAGGATCTTTGTG 568
QY 79 TGGCAAAACACACTTTTGGCTCAATTTATTCAAGCATCTGGCAAGAGCAAGCCCAACC 138
DB 569 TGGCAAAACACACTTTTGGCTCAATTTATTCAAGCATCTGGCAAGAGCAAGCCCAACC 628
QY 139 AGAACCCTCTTCCTCCCATGGAGCATCTGTCACCATGGCCATGGCTTACATGGGCT 198
DB 629 AGAACCCTCTTCCTCCCATGGAGCATCTGTCACCATGGCCATGGCTTACATGGGCT 688
QY 199 CCAGGGGCGAGCCGAGAGACAGATGGCCAAAGGTGCTTCAAGTTAAATGAAGTGGAGCCA 258
DB 689 CCAGGGGCGAGCCGAGAGACAGATGGCCAAAGGTGCTTCAAGTTAAATGAAGTGGAGCCA 748
QY 259 ATGCAGTTAACCCCATGACTCCAGAGAACTTTACAGCTGTGGTTCATGACAGCATCC 318
DB 749 ATGCAGTTAACCCCATGACTCCAGAGAACTTTACAGCTGTGGTTCATGACAGCATCC 808
QY 319 AGAAGGGTAGTTATCTCTGATGCGATTTTGCAGGCACAAAGTGCAGATAAAATCCATTCA 378
DB 809 AGAAGGGTAGTTATCTCTGATGCGATTTTGCAGGCACAAAGTGCAGATAAAATCCATTCA 868
QY 379 CCTTCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGAATTTATTACTTGGAAAGTG 438
DB 869 CCTTCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGAATTTATTACTTGGAAAGTG 928
QY 439 TCAATAAGCTGTTTGGTGAAGTCTGCGAGCTTCCGGGAGAAATATATTCAGCTCTGTGC 498
DB 929 TCAATAAGCTGTTTGGTGAAGTCTGCGAGCTTCCGGGAGAAATATATTCAGCTCTGTGC 988
|||

499 AGAATATTACTCTCAGAACCCAGGCGAGTAGACTTCTAGAAATGTGCGAGAGAGCTA 558
 989 AGAATATTACTCTCAGAACCCAGGCGAGTAGACTTCTAGAAATGTGCGAGAGAGCTA 1048
 559 GAAAAAGATTAATTCCTGGGTCAAGACTCAAAACAAAGGCAAAATCCCAAACTTGTAC 618
 1049 GAAAAAGATTAATTCCTGGGTCAAGACTCAAAACAAAGGCAAAATCCCAAACTTGTAC 1108
 619 CTGAAGGTTCTGTAGATGGGATACAGAGTGGTCTCTGTGAATGTCTACTTCCAAAG 678
 1109 CTGAAGGTTCTGTAGATGGGATACAGAGTGGTCTCTGTGAATGTCTACTTCCAAAG 1168
 679 GAAAGTGGAAATCCATTTGAGAGAACTAAATGGGCTTTATCCCTTCCGTGTAACCT 738
 1169 GAAAGTGGAAATCCATTTGAGAGAACTAAATGGGCTTTATCCCTTCCGTGTAACCT 1228
 739 CGGCTCAGCGCACACCTCTCAGAGTGTACTTGTGTAAGGCTAAACATTTGGATACA 798
 1229 CGGCTCAGCGCACACCTCTCAGAGTGTACTTGTGTAAGGCTAAACATTTGGATACA 1288
 799 TAGAAGACCTAAAGGCTCAGATTTAGAACTCCCATATGTCTGGAGATGTTAGCATGTTCT 858
 1289 TAGAAGACCTAAAGGCTCAGATTTAGAACTCCCATATGTCTGGAGATGTTAGCATGTTCT 1348
 859 TGTGCTTCCAGATGAATTCGCGATGTGTCCACTGGCTTGGAGTGTCTGGAAAGTGAA 918
 1349 TGTGCTTCCAGATGAATTCGCGATGTGTCCACTGGCTTGGAGTGTCTGGAAAGTGAA 1408
 919 TAACCTATGACAACTCAACAGTGGACAGCAAGAGCAAAATGGCTGAAGATGAAGTTG 978
 1409 TAACCTATGACAACTCAACAGTGGACAGCAAGAGCAAAATGGCTGAAGATGAAGTTG 1468
 979 AGGTATACATACCCAGTTCATTAATAGAGAGCAATTAATCACTCAGATCCATTCGAA 1038
 1469 AGGTATACATACCCAGTTCATTAATAGAGAGCAATTAATCACTCAGATCCATTCGAA 1528
 1039 GCATGGGATGGAGACGCTTCAACAGGAGCGGCGCAATTTCTCAGGATGTCGAGA 1098
 1529 GCATGGGATGGAGACGCTTCAACAGGAGCGGCGCAATTTCTCAGGATGTCGAGA 1588
 1099 GGAATGACCTGTTCTTCTGAAGTGTCCACCAAGCCTAGTGGATGTAATGAGGAGG 1158
 1589 GGAATGACCTGTTCTTCTGAAGTGTCCACCAAGCCTAGTGGATGTAATGAGGAGG 1648
 1159 GCATGAAGACGCGCTGGCAGAGAGTGTATGACAGGAGAACTGGACATGGAGGC 1218
 1649 GCATGAAGACGCGCTGGCAGAGAGTGTATGACAGGAGAACTGGACATGGAGGC 1708
 1219 CACAGTTTGTGGCAGATCATCGCTTTCTTTTCTTTTCTTTTATGATAGATAACCAAGTGCA 1278
 1709 CACAGTTTGTGGCAGATCATCGCTTTCTTTTCTTTTCTTTTATGATAGATAACCAAGTGCA 1768
 1279 TTTTATTTTGTGGGAGATTTCTCACCCTAAACCTAAGGCTGCTCTTTCGAAAAGAT 1338
 1769 TTTTATTTTGTGGGAGATTTCTCACCCTAAACCTAAGGCTGCTCTTTCGAAAAGAT 1828
 1339 TTTTGTAGATGAGCTGTGTGCTTCAGAAATGCTATTTCAAAATGCCCCAAATTTAGAGAT 1398
 1829 TTTTGTAGATGAGCTGTGTGCTTCAGAAATGCTATTTCAAAATGCCCCAAATTTAGAGAT 1888
 1399 GTTTTCTACATATTTCTGCTTCTTGAACAACTTCTGTACCCCTAAATATAAACAACAG 1458
 1889 GTTTTCTACATATTTCTGCTTCTTGAACAACTTCTGTACCCCTAAATATAAACAACAG 1948
 1459 AAATAATTAGACAAATGCTATTAATAACATGCAACCTTATTAATCATTTTGTCTTAA 1518
 1949 AAATAATTAGACAAATGCTATTAATAACATGCAACCTTATTAATCATTTTGTCTTAA 2008
 1519 AATGGATCATGCCCAATTTAGATTTTCTTACTATCATAGTTATTTTATACATTAACCTT 1578
 2009 AATGGATCATGCCCAATTTAGATTTTCTTACTATCATAGTTATTTTATACATTAACCTT 2068

1579 TTACTTTGTTATTATTATTTATATAATGTTAGTGTGTTTAAATTTATTTGCTCACTGCCTA 1638
 2069 TTACTTTGTTATTATTATTATATAATGTTAGTGTGTTTAAATTTATTTGCTCACTGCCTA 2128
 1639 TTTAATGTAGCTAATAAGTTATAGAACAGATGATCTGTTAATTTTCCTATCTAATAAAT 1698
 2129 TTTAATGTAGCTAATAAGTTATAGAACAGATGATCTGTTAATTTTCCTATCTAATAAAT 2188
 1699 GCGTTTAATTTCTCATATAATGAAGATAAGTAGGTATCCCTCCATGCCCTTCTGTAATA 1758
 2189 GCGTTTAATTTCTCATATAATGAAGATAAGTAGGTATCCCTCCATGCCCTTCTGTAATA 2248
 1759 AATATCTGGAAAAACATTAACATATAGCAATATATGTTATGTCATTTCTAGAAATA 1818
 2249 AATATCTGGAAAAACATTAACATATAGCAATATATGTTATGTCATTTCTAGAAATA 2308
 1819 CATTAACATATATATGTTCTGTTATCTTATATTTCAATTTCAAGTATATAATAAATAAACCCT 1878
 2309 CATTAACATATATATGTTCTGTTATCTTATATTTCAATTTCAAGTATATAATAAATAAACCCT 2368
 1879 GCTTCCAAACCAAAAAAATAAAAAA 1906
 2369 GCTTCCAAACCAAAAAAATAAAAAA 2396

RESULT 4
 AR380643 1884 bp DNA linear PAT 18-DEC-2003
 LOCUS Sequence 1188 from patent US 6607879.
 ACCESSION AR380643
 VERSION AR380643.1 GI:40089277
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 1884)
 AUTHORS Cocks B.G., Stuart S.G. and Seilhamer J.J.
 TITLE Compositions for the detection of blood cell and immunological response gene expression
 JOURNAL Patent: US 6607879-A 1188 19-AUG-2003;
 FEATURES Location/Qualifiers
 1. 1884
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 98.4%; Score 1876.6; DB 6; Length 1884;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1879; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

9 TTGCCCGTCAGACAGCAACTCAGAGAATAAACCAAGAGCAACCAAGATTTGAAACAATGGAG 68
 2 TTACCCGTCAGACAGCAACTCAGAGAATAAACCAAGAGCAACCAAGATTTGAAACAATGGAG 61
 69 GATCTTTGTGGGCAACACACACTTTGCCCTCAATTTTCAAGCATCTGGCAAAAAGCA 128
 62 GATCTTTGTGGGCAACACACACTTTGCCCTCAATTTTCAAGCATCTGGCAAAAAGCA 121
 129 AGCCCCACCCAGAACCTCTTCTCTCCCATGAGCATCTCGTCCACCATGGCCATGGTC 188
 122 AGCCCCACCCAGAACCTCTTCTCTCTCCCATGAGCATCTCGTCCACCATGGCCATGGTC 181
 189 TACATGGGCTCCAGGGCAGCACCCAGAGCAGATGGCAAGGTGCTTCAAGTTTAATGAA 248
 182 TACATGGGCTCCAGGGCAGCACCCAGAGCAGATGGCAAGGTGCTTCAAGTTTAATGAA 241
 249 GTGGAGCCAATGCAAGTTACCCCATGATCCAGAGAACTTTTACAGCTGTGGTTGATG 308
 242 GTGGAGCCAATGCAAGTTACCCCATGATCCAGAGAACTTTTACAGCTGTGGTTGATG 301
 309 CAGCAGATCCAGAGGGTAGTTTCTGATGGATTTTTCAGGCAACAAGCTGCAGATAAA 368
 302 CAGCAGATCCAGAGGGTAGTTTCTGATGGATTTTTCAGGCAACAAGCTGCAGATAAA 361

369 ATCCATTGATGCTTCCGCTCTCTAGCTCTGCAATCAATGCAATCCACAGGGAATATTATTA 428
 362 ATCCATTGATGCTTCCGCTCTCTAGCTCTGCAATCAATGCAATCCACAGGGAATATTATTA 421
 429 CTGGAAGTGTCATTAAGCTGTTTGGTGAGAGCTCTGCGAGCTTCCGGGGAAGATATATT 488
 422 CTGGAAGTGTCATTAAGCTGTTTGGTGAGAGCTCTGCGAGCTTCCGGGGAAGATATATT 481
 489 CGACTCTGTGAGAAATATTACTCTCAGAACCCAGGAGTACAGTCTCAGATGTCGA 548
 482 CGACTCTGTGAGAAATATTACTCTCAGAACCCAGGAGTACAGTCTCAGATGTCGA 541
 549 GAAGAAGCTGAGAAAGAAAGTAAATCTCTGGGTCAAGACTCAAAAGGCAAAATCCCA 608
 542 GAAGAAGCTGAGAAAGAAAGTAAATCTCTGGGTCAAGACTCAAAAGGCAAAATCCCA 601
 609 AACTTGTACCTGAAGTTCTGTAGATGGGATACAGAGTGTCTGGTGAATGCTGTC 668
 602 AACTTGTACCTGAAGTTCTGTAGATGGGATACAGAGTGTCTGGTGAATGCTGTC 661
 669 TACTTCAAGGAAAGTGGAAACTCCATTTGAGAGAACTAAATGGGCTTTATCTCTTC 728
 662 TACTTCAAGGAAAGTGGAAACTCCATTTGAGAGAACTAAATGGGCTTTATCTCTTC 721
 729 COTGTAACTCGGCTCAGCGCACACCTGTACAGATGATGCTGCTGGAAGCTAAAC 788
 722 COTGTAACTCGGCTCAGCGCACACCTGTACAGATGATGCTGCTGGAAGCTAAAC 781
 789 ATTGGATACATAGAACACCTAAAGGCTCAGATCTAGAACTCCCATATGCTGGAGATGTT 848
 782 ATTGGATACATAGAACACCTAAAGGCTCAGATCTAGAACTCCCATATGCTGGAGATGTT 841
 849 AGCATGTTCTTGTGTTCCAGATGAAATTTGCCGATGTCTCCACTGCTGGAGCTGCTG 908
 842 AGCATGTTCTTGTGTTCCAGATGAAATTTGCCGATGTCTCCACTGCTGGAGCTGCTG 901
 909 GAAAGTGAATTAACCTATGACAACTCAACAAAGTGGACCAAGGCAAAATGGCTGAA 968
 902 GAAAGTGAATTAACCTATGACAACTCAACAAAGTGGACCAAGGCAAAATGGCTGAA 961
 969 GATGAAGTTGAGGTATACATACCCAGTTCAAAATAGAGAGCAATTAAGTCAAGTCC 1028
 962 GATGAAGTTGAGGTATACATACCCAGTTCAAAATAGAGAGCAATTAAGTCAAGTCC 1021
 1029 ATTCTGAAAGAGCATGGGATGAGGAGCGCTTCAACAAAGGAGCGGCAATTTCTCAGGG 1088
 1022 ATTCTGAGAGCATGGGATGAGGAGCGCTTCAACAAAGGAGCGGCAATTTCTCAGGG 1081
 1089 ATGTCGGAGAGGATGACCTGTTCTTCTGAGTGTTCACCAAGCCATGCTGGATGTG 1148
 1082 ATGTCGGAGAGGATGACCTGTTCTTCTGAGTGTTCACCAAGCCATGCTGGATGTG 1141
 1149 AATGAGAGGAGGAGCATGAGGAGCGCTGGACAGGAGGTGTATGACAGGAGGAGTCCGA 1208
 1142 AATGAGAGGAGGAGCATGAGGAGCGCTGGACAGGAGGTGTATGACAGGAGGAGTCCGA 1201
 1209 CATGAGAGGAGGAGCATGAGGAGCGCTGGACAGGAGGTGTATGACAGGAGGAGTCCGA 1268
 1202 CATGAGAGGAGGAGCATGAGGAGCGCTGGACAGGAGGTGTATGACAGGAGGAGTCCGA 1261
 1269 ACCAAGTGCAATTTATTTTTCGGGAGATTTTTCCTCACCTCAAACTAAGCGTGTCTTC 1328
 1262 ACCAAGTGCAATTTATTTTTCGGGAGATTTTTCCTCACCTCAAACTAAGCGTGTCTTC 1321
 1329 TGCAAAAGATTTTGTAGATGAGTGTGCTGCTGAGATTTGCTTATTTCAAAATGTCGAAA 1388
 1322 TGCAAAAGATTTTGTAGATGAGTGTGCTGCTGAGATTTGCTTATTTCAAAATGTCGAAA 1381
 1389 ATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAATTTCTGCTACCCACTAAAT 1448
 1382 ATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAATTTCTGCTACCCACTAAAT 1441

1449 AABAACACAGAAATAATTAGCAATTTGCTATTATAAATGCAATGCAATCCATTAATCAATTT 1508
 1442 AABAACACAGAAATAATTAGCAATTTGCTATTATAAATGCAATGCAATCCATTAATCAATTT 1501
 1509 GGTCTTCTAAATAGGATCATGCCCATTTAGATTTTCTTACTATCATCAGTTATTTTATA 1568
 1502 GGTCTTCTAAATAGGATCATGCCCATTTAGATTTTCTTACTATCATCAGTTATTTTATA 1561
 1569 ACATTAACTTTTACTTTTGTATTATTATTATATAAATGCTGAGTTTAAATTTATTCG 1628
 1562 ACATTAACTTTTACTTTTGTATTATTATTATATAAATGCTGAGTTTAAATTTATTCG 1621
 1629 TCATGCTTATTAATGATAGTAAAGTTATAGAGGAGATGATCTGTTAAATTTCTTA 1688
 1622 TCATGCTTATTAATGATAGTAAAGTTATAGAGGAGATGATCTGTTAAATTTCTTA 1681
 1689 TCTAATAAATGCTTTAAATTTGTTTCTCATTAATGAAGAAATAGTAGGTATCCCTCCATGCC 1748
 1682 TCTAATAAATGCTTTAAATTTGTTTCTCATTAATGAAGAAATAGTAGGTATCCCTCCATGCC 1741
 1749 TCTGTAATAAATCTGGAAGAAATTAACAATAGCAATATATATGTTATGTCGATT 1808
 1742 TCTGTAATAAATCTGGAAGAAATTAACAATAGCAATATATATGTTATGTCGATT 1801
 1809 TCTGAAATACATACATATATATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTAT 1869
 1802 TCTGAAATACATACATATATATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTAT 1861
 1869 AAATAAATGCTTCTTCCAAACAAC 1891
 1862 AAATAAATGCTTCTTCCAAACAAC 1884

RESULT 5

HUMPA12 1884 bp mRNA linear PRI 27-APR-1993
 Human placental plasminogen activator inhibitor mRNA, complete cds.

LOCUS J02685.1 GI:189544
 DEFINITION plasminogen activator inhibitor; serine protease inhibitor.
 ACCESSION J02685.1
 VERSION
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM

REFERENCE
 1 (bases 1 to 1884)
 Ye, R.D., Wun, T.C. and Sadler, J.E.
 cDNA cloning and expression in Escherichia coli of a plasminogen
 activator inhibitor from human placenta
 J. Biol. Chem. 262 (8), 3718-3725 (1987)

JOURNAL MEDLINE
 PUBMED
 COMMENT

Original source text: Human placenta, cDNA to mRNA, clone
 lambda-PAI-75.1.
 Draft entry and computer-readable sequence for [1] kindly provided
 by R.D.Ye, 22-JAN-1987.

FEATURES
 Location/Qualifiers

1..1884
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 <1..1884
 /product="PAI mRNA"
 56..1303
 /note="plasminogen activator inhibitor"

mRNA

CDS

/protein_id="AAA36413.1"
 /translation="MEDLCVANTLFAALFKHLAKASPTONLFLSPWISSTMMVMYV
 GREGSTDMKVLQNEVCANAVTPNFTSCGEMQIQKQKYSVPAILQAOAADK
 IHSFRLSLASINASTDYLESVKNLFGKSAFSEYELCKQKYSVPAILQAOAADK
 CAEARKNSWVKTQTKIPLNLLPEGSVDGTRMVLVNAVYFGKWKTPREKING
 LYPFRVNSQRTQVMYLRKLNIGYIDELKQIILELPYAGDVSMFLLLPEIDAVS
 TGLELLESEITYDKLNKNTSKDMAEDEVYIIPQFLKEHYELRSILRSMGWDAPN

RIGIN		355 bp upstream of PstI site.	KGRANPSGMSERNDLFLSEVTHQAMVDVNEEGTEAAAGTGVMGRTGHHGGPQFVADH PFUFLMHKUTKCLIFFGRPCSP	
Query Match		98.4%; Score 1876.6; DB 9; Length 1884;		
Best Local Similarity		99.8%; Pred. No. 0;		
Matches 1879; Conservative.		0; Mismatches 4; Indels 0; Gaps 0;		
1	9	TTGCCGCTCAGACAGCAACTCAGAGATAACACAGAGAACACACCAATGGAACAATGGAG	68	
2	2	TTACCCGCTCAGACAGCAACTCAGAGATAACACAGAGAACACCAATGGAACAATGGAG	61	
3	69	GATCTTTGTGTGGCAAAACACACTCTTTCCCTCAATTTATTCAAGCATCTGGCAAAAGCA	128	
4	62	GATCTTTGTGTGGCAAAACACACTCTTTCCCTCAATTTATTCAAGCATCTGGCAAAAGCA	121	
5	129	AGCCCAACCAACACCTTCTCTCCCATGGAGCATCTGTCACCATGCGCATGGTC	188	
6	122	AGCCCAACCAACACCTTCTCTCCCATGGAGCATCTGTCACCATGCGCATGGTC	181	
7	189	TACATGGGCTCCAGGGGAGCAGCAGACGAGATGGCCAAAGTGCTTCAGTTTAAATGAA	248	
8	182	TACATGGGCTCCAGGGGAGCAGCAGACGAGATGGCCAAAGTGCTTCAGTTTAAATGAA	241	
9	249	GTGGAGCCCAATGACCTTACCCCAATGACTCCAGAGAACTTTACAGCTGTGGTTTCATG	308	
10	242	GTGGAGCCCAATGACCTTACCCCAATGACTCCAGAGAACTTTACAGCTGTGGTTTCATG	301	
11	309	CAGCAGATCCAGAGAGGAGTATCTCTGATGGATTTTGCAGGACCAAGCTGCAGATAAA	368	
12	302	CAGCAGATCCAGAGAGGAGTATCTCTGATGGATTTTGCAGGACCAAGCTGCAGATAAA	361	
13	369	ATCCATATCTCTCCGCTCTCAGCTCTGCAATCAATGATCCACAGGAAATTTTAA	428	
14	362	ATCCATATCTCTCCGCTCTCAGCTCTGCAATCAATGATCCACAGGAAATTTTAA	421	
15	429	CTGGAAGTGTCAATAAGCTGTGTGGTGAAGTCTGCGAGCTTCGGGAAGAAATATTT	488	
16	422	CTGGAAGTGTCAATAAGCTGTGTGGTGAAGTCTGCGAGCTTCGGGAAGAAATATTT	481	
17	489	CGACTCTGAGAAATATTTCTCTCAGAACCCAGGAGTAGACTTCTAGNATGTGCA	548	
18	482	CGACTCTGAGAAATATTTCTCTCAGAACCCAGGAGTAGACTTCTAGNATGTGCA	541	
19	549	GAGAAAGTGTAGAAAGAAATTAATTTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCA	608	
20	542	GAGAAAGTGTAGAAAGAAATTAATTTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCA	601	
21	609	AACCTGTACCTGAAAGTCTGTAGATGGGATACAGGATGGTCTGTGTAATGTGTC	668	
22	602	AACCTGTGTACCTGAAAGTCTGTAGATGGGATACAGGATGGTCTGTGTAATGTGTC	661	
23	669	TACTTCAAGGAAAGTGGAAACTCCATTTGAGAGAAACTAAATGGGCTTTATCTCTTC	728	
24	662	TACTTCAAGGAAAGTGGAAACTCCATTTGAGAGAAACTAAATGGGCTTTATCTCTTC	721	
25	729	CGTGTAACTGGCTCAGGCAACCTGTACAGATGATGTACTGCTGTAAGGCTAAAC	798	
26	722	CGTGTAACTGGCTCAGGCAACCTGTACAGATGATGTACTGCTGTAAGGCTAAAC	781	
27	789	ATTGGATACATAGAGACCTTAAGGCTCAGATTTAGAACTCCCATATGCTGGAGATGTT	848	
28	782	ATTGGATACATAGAGACCTTAAGGCTCAGATTTAGAACTCCCATATGCTGGAGATGTT	841	
29	849	AGCATGTTCTGTGCTTCCAGATGAAATTTGCCGATGTCTCACTGGCTTGGAGCTGCTG	908	
30	842	AGCATGTTCTGTGCTTCCAGATGAAATTTGCCGATGTCTCACTGGCTTGGAGCTGCTG	901	
31	909	GAAAGTGAATTAACCTATGACAACTCAACAAGTGGACCAAGCAAAAGTGGCTGAA	968	
32	902	GAAAGTGAATTAACCTATGACAACTCAACAAGTGGACCAAGCAAAAGTGGCTGAA	961	
33	969	GATGAAGTTGAGGTATACATACCCCAAGTTCAAATTTAGAGAGCATTTAGAACTCAGATCC	1028	

Db	962	GATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAGAGCATTAATGAATCAGATCC	1021	
Qy	1029	ATTCTGAAAAAGCATGGCGATGGAGAGCGCTTCAACAAGGAGCGGCCAATTTCTCAGGG	1088	
Db	1022	ATTCTGAAAAAGCATGGCGATGGAGAGCGCTTCAACAAGGAGCGGCCAATTTCTCAGGG	1081	
Qy	1089	ATGTCGGAGAGGAATGACCTGTTTCTTCTGAAGTGTTCACCAAGCCATGTGTGATGTG	1148	
Db	1082	ATGTCGGAGAGGAATGACCTGTTTCTTCTGAAGTGTTCACCAAGCCATGTGTGATGTG	1141	
Qy	1149	AATGAGAGGAGCACTGAAGCAGCGCTGGCAGAGAGTGTATGACAGGAGAACTGGA	1208	
Db	1142	AATGAGAGGAGCACTGAAGCAGCGCTGGCAGAGAGTGTATGACAGGAGAACTGGA	1201	
Qy	1209	CATGAGGCGCCACAGTTTGTGGCAGATCATCCGTTTCTTTTCTTTTATGATAAGATA	1268	
Db	1202	CATGAGGCGCCACAGTTTGTGGCAGATCATCCGTTTCTTTTCTTTTATGATAAGATA	1261	
Qy	1269	ACCAAGTGCAATTTATTTTTCGGCAGATTTTGTCTCACCTTAAACTTAAGCGTCTCTTC	1328	
Db	1262	ACCAAGTGCAATTTATTTTTCGGCAGATTTTGTCTCACCTTAAACTTAAGCGTCTCTTC	1321	
Qy	1329	TGCAAAAAGATTTTGTAGATGAGCTGTGTGCTCAGAAATTTGCTATTTCAAAATGCAAAA	1388	
Db	1322	TGCAAAAAGATTTTGTAGATGAGCTGTGTGCTCAGAAATTTGCTATTTCAAAATGCAAAA	1381	
Qy	1389	ATTGAGAGATGTTTCTACATATTTCTGCTCTCTGAAACAACTTCTGCTACCCACTAAT	1448	
Db	1382	ATTGAGAGATGTTTCTACATATTTCTGCTCTCTGAAACAACTTCTGCTACCCACTAAT	1441	
Qy	1449	AAAAACACAGAAATAATTAGACAAATTTGCTATTTATAACATGACAAACCTTAAATCATTT	1508	
Db	1442	AAAAACACAGAAATAATTAGACAAATTTGCTATTTATAACATGACAAACCTTAAATCATTT	1501	
Qy	1509	GGTCTTCTAAAAATGGGATCATGCCATTTAGATTTTCTTACTACAGTTTATTTTATA	1568	
Db	1502	GGTCTTCTAAAAATGGGATCATGCCATTTAGATTTTCTTACTACAGTTTATTTTATA	1561	
Qy	1569	ACATTAATCTTTTACTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	1628	
Db	1562	ACATTAATCTTTTACTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	1621	
Qy	1629	TCACTGCTATTTAATGTAAGTATAAAGTTATAGAAGCAGATGATCTGTTAATTTCTTA	1688	
Db	1622	TCACTGCTATTTAATGTAAGTATAAAGTTATAGAAGCAGATGATCTGTTAATTTCTTA	1681	
Qy	1689	TCTAATAATCCCTTTAATTTGTTCTCATAATGAAGAATAAGTAGGTATCCCTCCATGCC	1748	
Db	1682	TCTAATAATCCCTTTAATTTGTTCTCATAATGAAGAATAAGTAGGTATCCCTCCATGCC	1741	
Qy	1749	TTCTGTAATAAATATCTGGAATAAAACATTAACAATAGGCAATATATGTTATGTCAT	1808	
Db	1742	TTCTGTAATAAATATCTGGAATAAAACATTAACAATAGGCAATATATGTTATGTCAT	1801	
Qy	1809	TCTAGAATAACATACACATATATGTTCTGATCTTATTTCAATTTGCAAGTATATAAT	1868	
Db	1802	TCTAGAATAACATACACATATATGTTCTGATCTTATTTCAATTTGCAAGTATATAAT	1861	
Qy	1869	AAATAAACCTGCTTCCAAACAAAC 1891		
Db	1862	AAATAAACCTGCTTCCAAACAAAC 1884		

RESULT 6
 HSPA12R HSPA12R 1900 bp mRNA linear PRI 31-MAR-1995
 LOCUS Human mRNA for Arg-Serpin (plasminogen activator-inhibitor 2,
 DEFINITION PAI-2).
 ACCESSION Y00630
 VERSION Y00630.1 GI:35267
 KEYWORDS anti-urokinase; plasminogen activator-inhibitor type 2.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Webb, A.C., Collins, K.L., Snyder, S.E., Alexander, S.J.,
Rosenwasser, L.J., Eddy, R.L., Shows, T.B. and Auron, P.E.
TITLE Human monocyte Aq-Serpin cDNA: Sequence, chromosomal assignment,
and homology to plasminogen activator-inhibitor
J. Exp. Med. 166 (1), 77-94 (1987)

REFERENCE 2 (bases 1 to 1900)
AUTHORS Webb, A.C.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1987) Andrew C. Webb, Department of Biological
Sciences, Wellesley College, Wellesley, MA 02181, USA
COMMENT *source=LPS-stimulated monocytes; clone=pcd-1214
PAI-2 is a member CC of the serine protease inhibitor (serpin)
superfamily. It inhibits urokinase-type plasminogen activator. The
monocyte derived PAI-2 is distinct from the endothelial
cell-derived PAI-1.

FEATURES
source Location/Qualifiers
1..1900
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
73..1320
/note="unnamed protein product; PAI-2 precursor (AA -22 to
393)"
/codon_start=1
/protein_id="CAA68666.1"
/db_xref="GI:35268"
/db_xref="GOA:P05120"
/db_xref="SWISS-PROT:P05120"
/translation="MEDLCVANTILFALNLFKHLKASPTONLFLSPHSISSTMAVVM
GSRGSTDQAKVLTQNEGVANAVTMTPEFSCGFMQIQKGYSDAILQQAADK
IHSFSLASNAVSTNYLLESVNTKIPNLLPEGSVDGDMVNVAVPKGKWTFFPKKANG
CAEARKKINSWYKTKGIPNLLPEGSVDGDMVNVAVPKGKWTFFPKKANG
LVFVNSAORTVQVMYKREKLNIYIEDLKAQILELPVAGDVSMFLLLPDEIADVS
TGLELESITVYDKLANKWTSKDMAEDEVYIPOFKLEHYEURLSILRSMGMDAFN
KGRNPSGMSERNDLFLSEVFOHMDVNEEGTEAAAGTGGVMTGRTGHGGPQFVADH
PLFLIMHKITNCILFGRFSSP"
73..138
/notes="signal peptide (AA -22 to -1)"
139..1317
/product="PAI-2 (AA 1 - 393)"
295..297
/note="N-glycosylation site"
415..417
/note="N-glycosylation site"
1087..1089
/note="N-glycosylation site"
1210..1212
/note="active site of PAI-2"
1568..1577
/note="AT-rich sequence (pot. RNase target)"
1597..1609
/note="AT-rich sequence (pot. RNase target)"
1647..1703
/note="AT-rich sequence (pot. RNase target)"

sig_peptide
mat_peptide
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

Query Match 98.1%; Score 1872.6; DB 9; Length 1900;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1886; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

ORIGIN
1 GAGGACATTGCCCGTCAGACAGCAACTCAGAGATACACGAGAACCAACGATTGAAA 60
11 GAGGACATTGCCCGTCAGACAGCAACTCAGAGATACACGAGAACCAACGATTGAAA 70
61 CAATGGAGGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTATTCAAGCATCTGG 120
71 CAATGGAGGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTATTATTCAAGCATCTGG 130

121 CAAAGCAGAGCCCGCCAGAACCTTCTTCTCCATCCATGGAGCATCTCTCCACCATGG 180
131 CAAAGCAGAGCCCGCCAGAACCTTCTTCTCCATCCATGGAGCATCTCTCCACCATGG 190
181 CCATGGTCTACATGGGCTCCAGGGGAGCAGACCCAGAGACAGATGGCCAAAGTGCTTCAGT 240
191 CCATGGTCTACATGGGCTCCAGGGGAGCAGACCCAGAGACAGATGGCCAAAGTGCTTCAGT 250
241 TTAATGAAGTGGAGGCAATGAGTTACCCCAATGACTCCAGAGAACTTTCCAGAGTGTG 300
251 TTAATGAAGTGGAGGCAATGAGTTACCCCAATGACTCCAGAGAACTTTCCAGAGTGTG 310
301 GGTTCATGAGCAGATCCAGAGGGTGTATCTCTGATGCGATTTTCAGAGCAACAAGTG 360
311 GGTTCATGAGCAGATCCAGAGGGTGTATCTCTGATGCGATTTTCAGAGCAACAAGTG 370
361 CAGATAAATCCATCATCTCTCCGCTCTCTAGCTCTGCAATCAATGATCCACAGGGA 420
371 CAGATAAATCCATCATCTCTCCGCTCTCTAGCTCTGCAATCAATGATCCACAGGGA 430
421 ATTATTTACTGAAAGTGTCAATPAAGCTGTTGGTGAGAAAGTGTGCGAGCTTCCGGGAAG 480
431 ATTATTTACTGAAAGTGTCAATPAAGCTGTTGGTGAGAAAGTGTGCGAGCTTCCGGGAAG 490
481 ATATATTCGACTCTGTGAGAAATATTACTCTCAGAACCCAGCAGTAGACTTCTCTAG 540
491 ATATATTCGACTCTGTGAGAAATATTACTCTCAGAACCCAGCAGTAGACTTCTCTAG 550
541 AATGTGAGAGAGCTAGAAAAGATTAAATTCCTGGGTCAAGACTCAAAACCAAGGCA 600
551 AATGTGAGAGAGCTAGAAAAGATTAAATTCCTGGGTCAAGACTCAAAACCAAGGCA 610
601 AATCCCAACTGTTTACCTGAGAGTGTCTAGTAGGGGATACCAGAGTGTCTCTGGTGA 660
611 AATCCCAACTGTTTACCTGAGAGTGTCTAGTAGGGGATACCAGAGTGTCTCTGGTGA 670
661 ATGCTGTCTACTTAAAGAAAGTGAATAACTCCATTGAGAAACAAATAATGGCTTT 720
671 ATGCTGTCTACTTAAAGAAAGTGAATAACTCCATTGAGAAACAAATAATGGCTTT 730
721 ATCTCTTCCGTGTAAACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGGTGAAA 780
731 ATCTCTTCCGTGTAAACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGGTGAAA 790
781 AGCTAAACATGGATACATAGAACCTTAAGGCTCAGATCTTAGAACTCCCATATGCTG 840
791 AGCTAAACATGGATACATAGAACCTTAAGGCTCAGATCTTAGAACTCCCATATGCTG 850
841 GAGATGTTAGCATGTTCTTGTGCTTCCAGATGAAATTTGCCGATGTGTCCACTGGCTTG 900
851 GAGATGTTAGCATGTTCTTGTGCTTCCAGATGAAATTTGCCGATGTGTCCACTGGCTTG 910
901 AGCTGCTGAAAGTGAATAACTCTATGACAACTCAACAGTGGACCCAGCAAGCAAAA 960
911 AGCTGCTGAAAGTGAATAACTCTATGACAACTCAACAGTGGACCCAGCAAGCAAAA 970
961 TGGCTGAAAGTGAAGTTGAGGTATACATACCCCAAGTTCAAAATTAGAAGAGCATTTGAAC 1020
971 TGGCTGAAAGTGAAGTTGAGGTATACATACCCCAAGTTCAAAATTAGAAGAGCATTTGAAC 1030
1021 TCAGATCCATTTGAAAGCATGGGATGAGGAGCGCTTCAACAAAGGAGCGGCCCAATT 1080
1031 TCAGATCCATTTGAAAGCATGGGATGAGGAGCGCTTCAACAAAGGAGCGGCCCAATT 1090
1081 TCTCAGGGATGTCCGAGAGGAATGACCTGTTTCTTTCTGAAAGTGTTCACCAAGGCATGG 1140
1091 TCTCAGGGATGTCCGAGAGGAATGACCTGTTTCTTTCTGAAAGTGTTCACCAAGGCATGG 1150
1141 TGGATGTGAATGAGGAGGACCTGAACGCGCTGGCAGAGGAGTGTATGACAGGGA 1200
1151 TGGATGTGAATGAGGAGGACCTGAACGCGCTGGCAGAGGAGTGTATGACAGGGA 1210

```
1201 GAACGGACATGGAGGCCACAGATTGTGGCAGATCATCCGTTCTTTTCTTATATGCG 1260
1211 GAACGGACATGGAGGCCACAGATTGTGGCAGATCATCCCTTTCTTTCTTATATGCG 1270
1261 ATAAGTAACCAAGTGCATTTATTTTTCGGCAGATTTTGTCTCACCTAAACCTAAGCGT 1320
1271 ATAAGTAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCTAAACCTAAGCGT 1330
1321 GCTGCTCTCCAAAGATTTTGTAGAGTGTGCTGCTCAGATTCGCTATTTCAAT 1380
1331 GCTGCTCTCCAAAGATTTTGTAGAGTGTGCTGCTCAGATTCGCTATTTCAAT 1390
1381 TGGCAAAATTTAGAGATTTTCTACATATTTCTGCTCTTCTGCAACACTTCTGTACC 1440
1391 TGGCAAAATTTAGAGATTTTCTACATATTTCTGCTCTTCTGCAACACTTCTGTACC 1450
1441 CACTAATAAACAACAGATAATATAGACATTTCTATATACATGACACACCTATT 1500
1451 CACTAATAAACAACAGATAATATAGACATTTCTATATACATGACACACCTATT 1510
1501 AATCATTGCTCTCTAAATGGGATCATGCCAATTTAGATTTTCTTACTATCAGTTTA 1560
1511 AATCATTGCTCTCTAAATGGGATCATGCCAATTTAGATTTTCTTACTATCAGTTTA 1570
1561 TTTTATAACATTAACCTTTTACTTTGTTATTTATTTATTTATATGATGCTGAGTTTAA 1620
1571 TTTTATAACATTAACCTTTTACTTTGTTATTTATTTATTTATATGATGCTGAGTTTAA 1630
1621 ATTATTGCTCCTCCTCTATTTAATGTAGCTAATAAGTTATAGAGCAGATGATCTGTTA 1680
1631 ATTATTGCTCCTCCTCTATTTAATGTAGCTAATAAGTTATAGAGCAGATGATCTGTTA 1690
1681 ATTTCCTATCTAATAATGCTTTTAAATTTCTCATATTAATGAGATTAAGTATGCTTCC 1740
1691 ATTTCCTATCTAATAATGCTTTTAAATTTCTCATATTAATGAGATTAAGTATGCTTCC 1749
1741 CCATGCCCCTCTGTAATAATATCTGGAATAACATTAACAAATAGGCAAAATATATGTTA 1800
1750 CCATGCCCCTCTGTAATAATATCTGGAATAACATTAACAAATAGGCAAAATATATGTTA 1809
1801 TGTGCAATTTCTAGAAATACATACATATATATGCTGTATCTTATATTTCAATTTGCAAG 1860
1810 TGTGCAATTTCTAGAAATACATACATATATATGCTGTATCTTATATTTCAATTTGCAAG 1869
1861 TATATAATAATAAACCTGCTTCCAAACAAC 1891
1870 TATATAATAATAAACCTGCTTCCAAACAAC 1900

RESULT 7
HUMPAI2B 1880 bp mRNA linear PRI 07-JAN-1995
ACCU Human plasminogen activator inhibitor 2 (PAI-2) mRNA, complete cds.
DEFINITION
ACCESSION M18082
VERSION 1.0
KEYWORDS plasminogen activator inhibitor 2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1880)
AUTHORS Schleuning,W.D., Medcalf,R.L., Hession,C., Rothenbuhler,R., Shaw,A.
and Kruithof,E.K.
TITLE Plasminogen activator inhibitor 2: regulation of gene transcription
during phorbol ester-mediated differentiation of U-937 human
histiocyctic lymphoma cells
Mol. Cell. Biol. 7 (12), 4564-4567 (1987)
JOURNAL 88142852
MEDLINE 3325828
PUBMED
COMMENT Original source text: Human histiocyctic lymphoma cells, cDNA to
mRNA, clone pPAIJ.7.
FEATURES
Location/Qualifiers
1..1880
source
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="7q21.3-q22.1"
1..1880
/gene="PAI1"
/contig="PAI1"
32..1279
/gene="PAI1"
/contig="PAI1"
/feature="Plasminogen activator inhibitor type 2 precursor"
/translation="MEDLCVANTLFLALNLFKHLAKASPTQNLFLSPWSISSTWAMVYM
GSGSTDDQAKVLQFNEVGANAVTPTPEFTSCGFMOIQKSYDPAILOQAADK
IHHSFRSLGAINASTGYLLESLVKNLPEGSASPREYIRLCQKYSPPQAVDFLE
CASEARKINSWKTOTGKIPNLLPEGSVDGTRMVLNNAVYFKGKWTTPFKKNG
LTPRVNSAORTPVOMAVLREKLNATGYIEDLKAQJLELPYAGDVSMFLLPDLADVS
TGELLESEYTDKLNWTSKQNAEDVEVYIROPKLBHVELSILSMGMEDAFN
KGRANFGSMERNDFLSEVHQAVDNEEGTEAAGTGGVMTGRTGGGPPQFVADH
PFLFLIMHKITKCLIFGRFCSP"
32..97
/gene="PAI1"
/feature="Plasminogen activator inhibitor type 2 signal
peptide"
98..1276
/gene="PAI1"
/product="Plasminogen activator inhibitor type 2 protein"

ORIGIN
Query Match 96.9%; Score 1849.2; DB 9; Length 1880;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1865; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 34 AATAACAGAGAACCAACAGATTGAAACAATGGAGGATCTTTGTGGGCAAAACACACTCT 93
Db 3 ATTCCCAGAGAACCAACAGATTGAAACAATGGAGGATCTTTGTGGGCAAAACACACTCT 62
QY 94 TTGCCCTCAATTTTATCAAGCATCTGGCAAAAGCAAGCCCCACCCAGAACCTCTTCTCT 153
Db 63 TTGCCCTCAATTTTATCAAGCATCTGGCAAAAGCAAGCCCCACCCAGAACCTCTTCTCT 122
QY 154 CCCCATGAGCATCTGTCACCATCTGGCCATGCTTACATGGGCTCCAGGGGAGACACCG 213
Db 123 CCCATGAGCATCTGTCACCATCTGGCCATGCTTACATGGGCTCCAGGGGAGACACCG 182
QY 214 AAGACAGATGGCCAAAGTGTCTTCAAGTTAATGAAGTGGAGGCCAATGCAGTTACCCCA 273
Db 183 AAGACAGATGGCCAAAGTGTCTTCAAGTTAATGAAGTGGAGGCCAATGCAGTTACCCCA 242
QY 274 TGACTCCAGAGAACTTTTACAGCTGTGGGTTTCATGCAGCAGATCCAGAGGGTAGTTATC 333
Db 243 TGACTCCAGAGAACTTTTACAGCTGTGGGTTTCATGCAGCAGATCCAGAGGGTAGTTATC 302
QY 334 CTGATGCGATTTTCGAGGCAACAAGTGCAGATATAAATCCATTCATCTTCCGCTCTCTCA 393
Db 303 CTGATGCGATTTTCGAGGCAACAAGTGCAGATATAAATCCATTCATCTTCCGCTCTCTCA 362
QY 394 GCTCTGCAATCAATGTCATCCAGGGAATTTTACTGGAAGTGTCAATAGCTGTTTG 453
Db 363 GCTCTGCAATCAATGTCATCCAGGGGATTTTACTGGAAGTGTCAATAGCTGTTTG 422
QY 454 GTGAGAGTCTGCGAGCTTCCGGGAAGATATATTCTGACTGTGTCAAGATATATTCTCT 513
Db 423 GTGAGAGTCTGCGAGCTTCCGGGAAGATATATTCTGACTGTGTCAAGATATATTCTCT 482
QY 514 CAGAACCCCGAGGAGTACATCTCTAGAAATGTGCAAGAAAGCTAGAAAAAGATTAATT 573
Db 483 CAGAACCCCGAGGAGTACATCTCTAGAAATGTGCAAGAAAGCTAGAAAAAGATTAATT 542
QY 574 CTTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTGTGTAAGGTTCTGTAG 633
Db 543 CTTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTGTGTAAGGTTCTGTAG 602
```

634 ATGGGATACAGGATGGTCTGGTGAATGCTGTCTACTCTCAAGGAAAGTGGAAACTC 693
 Db
 603 ATGGGATACAGGATGGTCTGGTGAATGCTGTCTACTCTCAAGGAAAGTGGAAACTC 662
 2y
 694 CATTTGAGAAGAACTAAATGGGCTTTATCTCTTCCGTTGTAATCTCGCTCAGGCAAC 753
 Db
 663 CATTTGAGAAGAACTAAATGGGCTTTATCTCTTCCGTTGTAATCTCGCTCAGGCAAC 722
 2y
 754 CTGTACAGATGATGATCTTGGTGAAGAAAGTCAAACTTGGATACATAGAGACCTAAAGG 813
 Db
 723 CTGTACAGATGATGATCTTGGTGAAGAAAGTCAAACTTGGATACATAGAGACCTAAAGG 782
 2y
 814 CTCAGATTTAGAACTCCCATATGCTGGAGATGTTAGCATGTTCTTGTCTCCAGATG 873
 Db
 783 CTCAGATTTAGAACTCCCATATGCTGGAGATGTTAGCATGTTCTTGTCTCCAGATG 842
 2y
 874 AAATTTGCCATGCTGCTCACTGGCTTGGAGTCTGTGGAAGTGAATTAACCTATGACAAAC 933
 Db
 843 AAATTTGCCATGCTGCTCACTGGCTTGGAGTCTGTGGAAGTGAATTAACCTATGACAAAC 902
 2y
 934 TCAACAAGTGGACCAAGCAAAAGTCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 993
 Db
 903 TCAACAAGTGGACCAAGCAAAAGTCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 962
 2y
 994 AGTTCAAATTTAGAAAGCAATTTATGAATCAGATCCATTTCTGAAAAGCATGGCATGGAGG 1053
 Db
 963 AGTTCAAATTTAGAAAGCAATTTATGAATCAGATCCATTTCTGAAAAGCATGGCATGGAGG 1022
 2y
 1054 AGCCCTTCAACAAGGACGGCCAAATTTCTCAGGATGTCGGAGGAAGTGAAGTGAAGTGAAG 1113
 Db
 1023 AGCCCTTCAACAAGGACGGCCAAATTTCTCAGGATGTCGGAGGAAGTGAAGTGAAGTGAAG 1082
 2y
 1114 TTCTGAGTGTTCCTCAACCAAGCCATGCTGATGTAATGAGGAGGCACTGAAGCAGCCG 1173
 Db
 1083 TTCTGAGTGTTCCTCAACCAAGCCATGCTGATGTAATGAGGAGGCACTGAAGCAGCCG 1142
 2y
 1174 CTGGCAGAGGATGTTATGACAGGAGGATCTGACATGAGGAGGCACTGAAGCAGCCG 1233
 Db
 1143 CTGGCAGAGGATGTTATGACAGGAGGATCTGACATGAGGAGGCACTGAAGCAGCCG 1202
 2y
 1234 ATCATCCGTTCTTTTCTTATTTATGATGAATGATGATGATGATGATGATGATGATGATGAT 1293
 Db
 1203 ATCATCCGTTCTTTTCTTATTTATGATGAATGATGATGATGATGATGATGATGATGATGAT 1262
 2y
 1294 GATTTGCTCACCCTAAACTAGGCTGCTCTGCAAGATTTTGTAGATGATGATGATGATGATGAT 1353
 Db
 1263 GATTTGCTCACCCTAAACTAGGCTGCTCTGCAAGATTTTGTAGATGATGATGATGATGATGATGAT 1322
 2y
 1354 GTGTGCTCAGAAATGCTATTTCAAATTTGCAAAATTTAGAGATGTTTCTACATATTT 1413
 Db
 1323 GTGTGCTCAGAAATGCTATTTCAAATTTGCAAAATTTAGAGATGTTTCTACATATTT 1382
 2y
 1414 CTGCTCTTGAACAACTCTGCTACCCACTTAATTAATAAACAACAGAAATATTAGACAAT 1473
 Db
 1383 CTGCTCTTGAACAACTCTGCTACCCACTTAATTAATAAACAACAGAAATATTAGACAAT 1442
 2y
 1474 TGTCTATTAAACATCAACCCCTATTAACTTTGCTTTCTAAATTTGAGATGATGATGATGATGAT 1533
 Db
 1443 TGTCTATTAAACATCAACCCCTATTAACTTTGCTTTCTAAATTTGAGATGATGATGATGATGAT 1502
 2y
 1534 ATTTAGATTTCTTACTATCAGTTTATTTTAAACATTTTAACTTTTACTTTTGTATTTA 1593
 Db
 1503 ATTTAGATTTCTTACTATCAGTTTATTTTAAACATTTTAACTTTTACTTTTGTATTTA 1562
 2y
 1594 TTATTTTATAATAGTGGATTTTAAATTTATGCTCACTGCTCTATTTTAAATTTAGCTAAT 1653
 Db
 1563 TTATTTTATAATAGTGGATTTTAAATTTATGCTCACTGCTCTATTTTAAATTTAGCTAAT 1621
 2y
 1654 AAGTTTATGAAGCAGATGATCTGTTAATTTCTTATCTTAATAAATGCTTTTAAATTTGTTCT 1713
 Db
 1622 AAGTTTATGAAGCAGATGATCTGTTAATTTCTTATCTTAATAAATGCTTTTAAATTTGTTCT 1681
 2y
 1714 CATAATGAAGAAATAGTAGGTATCCCTCCATGCTCTCTGTAATAAATATCTGGAAAAA 1773

1682 CATATGAAGAAATAGTAGGTATCCCTCCATGCTCTTATATAAATATCTGGAAAAA 1741
 Db
 1774 CATTAACAATAGGCAAAATATATGTTATGCTGCAATTTCTAGAAATACATACACATATATA 1833
 2y
 1742 CATTAACAATAGGCAAAATATATGTTATGCTGCAATTTCTAGAAATACATACACATATATA 1801
 Db
 1834 TGTCTGTATCTTATATTCATATTCAGATGATATAATAATAAATCTGCTCCAAACAACAA 1893
 2y
 1802 TGTCTGTATCTTATATTCATATTCAGATGATATAATAATAAATCTGCTCCAAACAACAA 1861
 Db
 1894 AAAATATAAATAAAG 1907
 2y
 1862 AAAATATAAATAAAG 1875
 Db
 RESULT 8
 AR380644
 LOCUS AR380644 1854 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 1189 from patent US 6607879.
 ACCESSION AR380644
 VERSION AR380644.1 GI:40088278
 KEYWORDS
 SOURCE Unknown.
 ORGANISM
 Cocks, B.G., Stuart, S.G. and Seilhamer, J.J.
 TITLE Compositions for the detection of blood cell and immunological
 response gene expression
 JOURNAL Patent: US 6607879-A 1189 19-AUG-2003;
 FEATURES
 source
 1. 1854
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Query Match 96.8%; Score 1847.6; DB 6; Length 1854;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1850; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 15 GTCAGACAGCACTCAGAGAATAACCCAGAGCAACACAGATTGAAACATGAGGATCTT 74
 Db 1 GTCAGACAGCACTCAGAGAATAACCCAGAGCAACACAGATTGAAACATGAGGATCTT 60
 75 TGTGTGGCAACACACACTCTTTGCCCTCAATTTTCAAGCATCTGGCAAAAGCAAGCCCC 134
 61 TGTGTGGCAACACACACTCTTTGCCCTCAATTTTCAAGCATCTGGCAAAAGCAAGCCCC 120
 135 ACCAGAACTCTTCTCTCCCATGAGCATCTGCTCCACCATGGCCATGGTCTACATG 194
 Db 121 ACCAGAACTCTTCTCTCCCATGAGCATCTGCTCCACCATGGCCATGGTCTACATG 180
 195 GGTCCAGGGGAGCACCAGAGCAGATGGCCAAAGGTGCTTCAAGTTTAAATGAAGTGGGA 254
 Db 181 GGTCCAGGGGAGCACCAGAGCAGATGGCCAAAGGTGCTTCAAGTTTAAATGAAGTGGGA 240
 255 GCAATGAGTTACCCCATGACTCCAGAGAACTTTACCAGCTGTGGTTTATGAGCAG 314
 Db 241 GCAATGAGTTACCCCATGACTCCAGAGAACTTTACCAGCTGTGGTTTATGAGCAG 300
 315 ATCCAGAGGGTATCTCTGATGCGATTTTGAGGCAAGCTGCAGATAAATTCAT 374
 Db 301 ATCCAGAGGGTATCTCTGATGCGATTTTGAGGCAAGCTGCAGATAAATTCAT 360
 375 TCATCTTCGCTCTCTCAGCTCTGCAATCATCATGATCCAGAGAAATTTTACTGGAA 434
 Db 361 TCATCTTCGCTCTCTCAGCTCTGCAATCATCATGATCCAGAGAAATTTTACTGGAA 420
 435 AGTGTCAATAAGCTGTTTGGTGAAGTCTGCGAGCTTCCGGGAGAAATATATTCGACTC 494
 Db 421 AGTGTCAATAAGCTGTTTGGTGAAGTCTGCGAGCTTCCGGGAGAAATATATTCGACTC 480


```
495 TGTGAGAAATATTACTCTCTCAGAACCCCGAGCAGTAGACTTCTCTAGAAATGTGCAGAGAA 554
481 TGTGAGAAATATTACTCTCTCAGAACCCCGAGCAGTAGACTTCTCTAGAAATGTGCAGAGAA 540
555 GCTAGAAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTGT 614
541 GCTAGAAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTGT 600
615 TTACCTGAAGGTTCTGTAGATGGGGATACCAAGGATGGTCCCTGGTGAATGCTGTCTACTTC 674
601 TTACCTGAAGGTTCTGTAGATGGGGATACCAAGGATGGTCCCTGGTGAATGCTGTCTACTTC 660
675 AAGGAAAGTGAAGAACTCCATTTGAGAGAACTAAATGGGCTTATCTCTTCGGTGA 734
661 AAGGAAAGTGAAGAACTCCATTTGAGAGAACTAAATGGGCTTATCTCTTCGGTGA 720
735 AACTCGGCTCAGCGCACACCTGTGACAGATGATGTACTTCCGTGAAGGCTAAACATTTGA 794
721 AACTCGGCTCAGCGCACACCTGTGACAGATGATGTACTTCCGTGAAGGCTAAACATTTGA 780
795 TACATGAAGACCTAAAGGCTCAGATCTAGAACTCCCATATGCTGGAGATGTTAGCATG 854
781 TACATGAAGACCTAAAGGCTCAGATCTAGAACTCCCATATGCTGGAGATGTTAGCATG 840
855 TTCTGTGTCTCCAGATGAATTTCCCGATGTGTCCACTGGCTTGGAGCTGCTGGAAGT 914
841 TTCTGTGTCTCCAGATGAATTTCCCGATGTGTCCACTGGCTTGGAGCTGCTGGAAGT 900
915 GAATACCTATGACAACTCAACAAAGTGGACCAAGCAAGACAAAATGCTCAAGATGA 974
901 GAATACCTATGACAACTCAACAAAGTGGACCAAGCAAGACAAAATGCTCAAGATGA 960
975 GTTGAGGTATACATACCCAGTCAAAATAGAGAGCAATATGAATCAGATCCATTTCTG 1034
961 GTTGAGGTATACATACCCAGTCAAAATAGAGAGCAATATGAATCAGATCCATTTCTG 1020
1035 ARAAGCATGGCATGGAGGCGCTTCAACAGGACGGGCCAAATTTCTCAGGATGTCG 1094
1021 ARAAGCATGGCATGGAGGCGCTTCAACAGGACGGGCCAAATTTCTCAGGATGTCG 1080
1095 GAGAGGAATGACCTGTTTCTTGAAGTGTTCACCAAGCCATGCTGATGTAATGAG 1154
1081 GAGAGGAATGACCTGTTTCTTGAAGTGTTCACCAAGCCATGCTGATGTAATGAG 1140
1155 GAGGGCACTGAAGCAGCGCTGGCAGAGGTGTATGACAGGGAGAACTGCAATGGA 1214
1141 GAGGGCACTGAAGCAGCGCTGGCAGAGGTGTATGACAGGGAGAACTGCAATGGA 1200
1215 GGCCACAGTTTGTGGCAGATCATCCGTTCTTTCTTATTATGATGAATAAGCAAG 1274
1201 GGCCACAGTTTGTGGCAGATCATCCGTTCTTTCTTATTATGATGAATAAGCAAG 1260
1275 TGCAATTTATTTTTCGGCAGATTTTGTCCACCTAAACCTAAGCGTGTCTTCTGAAA 1334
1261 TGCAATTTATTTTTCGGCAGATTTTCTTCCACCTAAACCTAAGCGTGTCTTCTGAAA 1320
1335 AGATTTTGTAGATGAGCTGTGTGCTCAGAAATGCTATTTCAAAATGCGCAAAATTTAG 1394
1321 AGATTTTGTAGATGAGCTGTGTGCTCAGAAATGCTATTTCAAAATGCGCAAAATTTAG 1380
1395 AGATGTTTCTACATATTTCTGCTCTCTGAAACAACTTCTGCTACCCACTAAATAAAG 1454
1381 AGATGTTTCTACATATTTCTGCTCTCTGAAACAACTTCTGCTACCCACTAAATAAAG 1440
1455 ACAGAAATATAGACAAATGCTATTAATAACATGACAAACCTTAAATCATTTGCTT 1514
1441 ACAGAAATATAGACAAATGCTATTAATAACATGACAAACCTTAAATCATTTGCTT 1500
1515 CTAAGATGGGATCATGCCATTTAGATTTCTTACTATCAGTTTATTTTATAACATTA 1574
1501 CTAAGATGGGATCATGCCATTTAGATTTCTTACTATCAGTTTATTTTATAACATTA 1560
1575 ACTTTTACTTTGTTATTTATTTTATTAATAAATGGTGGATTTTAAATTTAGCTCACTG 1634
```

```
1561 ACATTACTTTGTTATTATTATTATTATTATTAATGCTGAGTTTAAATTTATGCTCACTG 1620
1635 CCTATTTAATGTAGCTAATAAGTTATAGAGCAGATGATCTGTTAATTTCCCTATCTAAT 1694
1621 CCTATTTAATGTAGCTAATAAGTTATAGAGCAGATGATCTGTTAATTTCCCTATCTAAT 1680
1695 AAATGCCCTTTAATTTGTTCTCATAATGAAGAATAAGTAGTATCCCTCCAGCCCTTCTGT 1754
1681 AAATGCCCTTTAATTTGTTCTCATAATGAAGAATAAGTAGTATCCCTCCAGCCCTTCTGT 1740
1755 AATAAATATCTCGAAGAAACATTAACCAATAGGCAATATATGTTATGTCATTTCTAGA 1814
1741 AATAAATATCTCGAAGAAACATTAACCAATAGGCAATATATGTTATGTCATTTCTAGA 1800
1815 AATACATAACACATATATATCTGTTCTATATTTCAATTCGAAGTATATAAT 1868
1801 AATACATAACACATATATATCTGTTCTATATTTCAATTCGAAGTATATAAT 1854

RESULT 9
LOCUS HUMPAI2A
DEFINITION Human plasminogen activator inhibitor mRNA, complete cds.
ACCESSION J03603
VERSION J03603.1 GI:189546
KEYWORDS plasminogen activator inhibitor.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1854)
AUTHORS Antalics,T.M., Clark,M.A., Barnes,T., Lehrbach,P.R., Devine,P.L.,
Schevzov,G., Goss,N.H., Stephens,R.W. and Tolstoshev,P.
TITLE Cloning and expression of a cDNA coding for a human
monocyte-derived plasminogen activator inhibitor
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (4), 985-989 (1988)
MEDLINE 88125032
PubMed 3257578
COMMENT Original source text: Human monocyte cell line U937, cDNA to mRNA,
clone BTA1445.
Draft entry and printed copy of sequence for (1) kindly provided by
T.M.Antalics, 01/19/88. Polyadenylation signals are located at
positions 1432-1437, 1637-1642, 1678-1683 and 1741-1746.

FEATURES
source
1..1854
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="1q23"
1..1854
/gene="FCER1A"
49..1296
/gene="FCER1A"
/notes="plasminogen activator inhibitor precursor"
/codon_start=1
/protein_id="AAA60004.1"
/db_xref="GI:189547"
/translat="MED:CVANTLFPALNLFKHLAKASPTQNLFLSPWSISSTMAVYM
GSRGSDQAKVLQFNEVGANAVTPTNPENFSGCFMQOIQKSYPDAILQAOADK
IHSSFRSLSSAINASTNYLLESVNKLFEKSKASPREYIRLCQKYSPQAVDFLE
CAEAPRVNSKINSWPTQTKIPNLLPEGSVDGTRMLVNVYFKGKWKTPFEKFLNG
LYPEPVNSAORTPVQMYLREKLIAGYIEDLKAQILLELPYAGDVSMELLPLDEIADYS
TGLELLSEITYDKLNKTKSDKXADSEVEVYIPQKLEHVELRSLRSMGMDAEN
KGRANFSGMSERNDLFLSEVFHQWVDVNEEGTEAAAGTGGVWTGTRGHGPGFVADH
PFPLIWHKINCLIFFGRFSSP"
52..1293
/mat_peptide
/gene="FCER1A"
/product="plasminogen activator inhibitor"
ORIGIN 13 bp upstream of DdeI site.

Query Match 96.8%; Score 1847.6; DB 9; Length 1854;
```

[illegible]


```

Db      1801 AATACATAACATATATATGCTGATCTTATTTCAATTGCAAGTATATAT 1854
|||||
RESULT 11
A16054      2409 bp      RNA      linear      PAT 23-MAR-1994
LOCUS      molecule with the activity of miniactivin.
DEFINITION
ACCESSION  A16054
VERSION     A16054.1 GI:512777
KEYWORDS   synthetic construct
SOURCE      synthetic construct
ORGANISM   artificial sequences.
AUTHORS     1 (bases 1 to 2409)
JOURNAL
FEATURES
    Patent: WO 8705628-A 11 24-SEP-1987;
    Location/Qualifiers
        1..2409
        /organism="synthetic construct"
        /mol_type="unassigned RNA"
        /db_xref="taxon:32830"
        49..1296
        /codon_table=1
        /transl_table=11
        /product="protein with miniactivin activity"
        /protein_id="CAA01261.1"
        /db_xref="GI:512778"
        /db_xref="REMBEMBL:CAA01261"
        /translation="MEDLVANTLPAINTFKHLAKASPTONLFLSPHSISSTWAMVYM
        GSRGSTDQMAKVLQNEVGANAVTPMTFENFTSCGFMQIQKSYFDALLOQAABX
        IHSFSLSSANASTGNLYLVSVNKLFGKSFASPREYIRLCOKYSSFPQVDFPLS
        CAEARKKINSVKTQTKGIPNLPEGSVDGTRVNLVNAVFKWKPFEKFLNG
        LYPFRVNSQRTFVQMYLREKLNIGYIEDLKAQIILEPYAGDVSMFLPLPDEIADVS
        TGLELLESIITVDKLNKNTSKDQWAEDEVYIPOFKLESHYELRSILRSGMEDAFN
        KGRANFSGMSEKNDLFLSVFHOAMVDVNEEGTEAAAGTGGVMTGRTGHGQPQVADH
        PLFLIMHKITNCILFFGRFSP"
    ORIGIN
        Query Match      96.8%; Score 1847.6; DB 6; Length 2409;
        Best Local Similarity 99.8%; Pred. No. 0;
        Matches 1850; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
        15  GTGAGACAGCAACTCAGAGAAATTAACAGAGAAACAACAGATTTGAAACAATGGAGATCTT 74
        Db      1  GTGAGACAGCAACTCAGAGAAATTAACAGAGAAACAACAGATTTGAAACAATGGAGATCTT 60
        75  TGTGTGGCAACACACTCTTTGCCCTCAATTTATTCAGCATCTGCGAAGAGCAAGCCCC 134
        Db      61  TGTGTGGCAACACACTCTTTGCCCTCAATTTATTCAGCATCTGCGAAGAGCAAGCCCC 120
        135  ACCGAGAACCTCTTCTCTCCCATGGAGCATCTGTCACCATGGCCCATGGCTCTACATG 194
        Db      121  ACCGAGAACCTCTTCTCTCCCATGGAGCATCTGTCACCATGGCCCATGGCTCTACATG 180
        195  GGTCTCAGGGGCGACCGGAGACCGATGCGCCAGGTGCTTCAGTTTAATGAAGTGGGA 254
        Db      181  GGTCTCAGGGGCGACCGGAGACCGATGCGCCAGGTGCTTCAGTTTAATGAAGTGGGA 240
        255  GCCAATGCGATTACCCCATGACTCCAGAGAACTTTTACCAGCTGTGGGTTTATGCAAGAG 314
        Db      241  GCCAATGCGATTACCCCATGACTCCAGAGAACTTTTACCAGCTGTGGGTTTATGCAAGAG 300
        315  ATCCAGAGGGTAGTTATCTGATGCGATTTTTCAGGACACAGCTGCGAGATAAATCCAT 374
        Db      301  ATCCAGAGGGTAGTTATCTGATGCGATTTTTCAGGACACAGCTGCGAGATAAATCCAT 360
        375  TCATCTTCGCTCTCTCAGCTGCGAATCAATGATCCAGGGAATTTTACTGGAA 434
        Db      361  TCATCTTCGCTCTCTCAGCTGCGAATCAATGATCCAGGGAATTTTACTGGAA 420
        435  AGTGTCAATAAGCTGTTTGGTGAAGTCTCGGAGCTTCCGGGAGAGATATATTGAGCTC 494
        |||||

```

```

Db      421  AGTGTCAATAAGCTGTTGGTGAGAAGTCTCGAGCTTCGGGAAGAATATATTTCGACTC 480
Qy      495  TGTGAGAATAATTACTCTCTCAGAACCCCGAGGAGTAGACTTCTCTAGAAATGTGCAGAGAA 554
Db      481  TGTGAGAATAATTACTCTCTCAGAACCCCGAGGAGTAGACTTCTCTAGAAATGTGCAGAGAA 540
Qy      555  GCTAGAAAAAGATTAAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAAATCCCAAATCTG 614
Db      541  GCTAGAAAAAGATTAAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAAATCCCAAATCTG 600
Qy      615  TTACCTGAAGGTTCTGTAGATGGGATACAGAGTGTCTCTGTGAATGTCTGTCTACTTTC 674
Db      601  TTACCTGAAGGTTCTGTAGATGGGATACAGAGTGTCTCTGTGAATGTCTGTCTACTTTC 660
Qy      675  AAAGGAAAGTGGAAAACTCCATTGAGAAGAACTAAATGGGCTTTATCTTCCGCTGTA 734
Db      661  AAAGGAAAGTGGAAAACTCCATTGAGAAGAACTAAATGGGCTTTATCTTCCGCTGTA 720
Qy      735  AACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGGCTGAAAAAGCTAAACATTGGA 794
Db      721  AACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGGCTGAAAAAGCTAAACATTGGA 780
Qy      795  TACATAGAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGATGTTAGCATG 854
Db      781  TACATAGAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGATGTTAGCATG 840
Qy      855  TTCTTTGTTGCTTCCAGATGAAATTTGCCGATGCTCTCACTCGCTTGGAGCTCTGGAAGT 914
Db      841  TTCTTTGTTGCTTCCAGATGAAATTTGCCGATGCTCTCACTCGCTTGGAGCTCTGGAAGT 900
Qy      915  GAAATTAACCTATGACAAACTCAACAAGTGGACAGCAAGCAAAATGGCTGAAGATGAA 974
Db      901  GAAATTAACCTATGACAAACTCAACAAGTGGACAGCAAGCAAAATGGCTGAAGATGAA 960
Qy      975  GTTGAAGTATACATACCCAGTTCAAAATTAGAGAGAGATTATGAACCTCAGATCCATTCTG 1034
Db      961  GTTGAAGTATACATACCCAGTTCAAAATTAGAGAGAGATTATGAACCTCAGATCCATTCTG 1020
Qy      1035  AAAAGCATGGCATGGAGGAGCGCTTCAACAAGGAGCGGCAATTTCTCAGGGAGTGTG 1094
Db      1021  AAGAGCATGGCATGGAGGAGCGCTTCAACAAGGAGCGGCAATTTCTCAGGGAGTGTG 1080
Qy      1095  GAGAGAAATGACTGTTCTTCTGAAAGTGTTCACCAAGCCATGGTGGATGTGATGAG 1154
Db      1081  GAGAGAAATGACTGTTCTTCTGAAAGTGTTCACCAAGCCATGGTGGATGTGATGAG 1140
Qy      1155  GAGGGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGGAGAACTGGACATGGA 1214
Db      1141  GAGGGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGGAGAACTGGACATGGA 1200
Qy      1215  GGGCCACAGTTTGTGCGAGATCATCGTTTCTTTTCTTTTCTTTTATGATGATAAGTACC 1274
Db      1201  GGGCCACAGTTTGTGCGAGATCATCGTTTCTTTTCTTTTCTTTTATGATGATAAGTACC 1260
Qy      1275  TGCATTTTATTTTTCGGCAGATTTTGTCTACCCCTAAAACCTAAGCGTGTCTTCTGCAAA 1334
Db      1261  TGCATTTTATTTTTCGGCAGATTTTGTCTACCCCTAAAACCTAAGCGTGTCTTCTGCAAA 1320
Qy      1335  AGATTTTCTGATGAGCTGTGTCCTCAGAAATGCTATTTCAAATTCGCAAAAATTTAG 1394
Db      1321  AGATTTTCTGATGAGCTGTGTCCTCAGAAATGCTATTTCAAATTCGCAAAAATTTAG 1380
Qy      1395  AGATGTTTTCTACATATTTCTGCTCTTCTGAAACAACTTCTGCTACCCCTAATAAATAAC 1454
Db      1381  AGATGTTTTCTACATATTTCTGCTCTTCTGAAACAACTTCTGCTACCCCTAATAAATAAC 1440
Qy      1455  ACAGAAATATTAGACAAATGCTATTATTAACATGACAAACCTATTATCAATTTGGTCTTT 1514
Db      1441  ACAGAAATATTAGACAAATGCTATTATTAACATGACAAACCTATTATCAATTTGGTCTTT 1500
Qy      1515  CTAATAATGGGATCATGCCCCATTTAGATTTTCTTACTATCAGTTTATTTTATACATTA 1574
Db      1501  CTAATAATGGGATCATGCCCCATTTAGATTTTCTTACTATCAGTTTATTTTATACATTA 1560

```

2y	1575	ACTTTTACTTTGTTATTTATTTATTTATATAAATGGTGAGTTTTTAAATATTATGCTCACTG	1634
Db	1561	ACTTTTACTTTGTTATTTATTTATTTATATAAATGGTGAGTTTTTAAATATTATGCTCACTG	1620
2y	1635	CCATTATTAATGTAGCTAATAAAGTTATAGAAGCAGATGATCTGTTAATTTCCCTATCTAAT	1694
Db	1621	CCATTATTAATGTAGCTAATAAAGTTATAGAAGCAGATGATCTGTTAATTTCCCTATCTAAT	1680
2y	1695	AAATGCGCTTTAAATGTCTCATTAATGAAGAAATGAATAGGTATCCCTCCATGCGCTTCTGT	1754
Db	1681	AAATGCGCTTTAAATGTCTCATTAATGAAGAAATGAATAGGTATCCCTCCATGCGCTTCTGT	1740
2y	1755	AATPAAATATCTGGAATAAACATTAAACAATAGGCAATATATGTTATGTGCATTTCTTAGA	1814
Db	1741	AATPAAATATCTGGAATAAACATTAAACAATAGGCAATATATGTTATGTGCATTTCTTAGA	1800
2y	1815	AATACATAACACATATATATGTCGTGATCTTATATTCAAATTCGAAGTATATAAT	1868
Db	1801	AATACATAACACATATATATGTCGTGATCTTATATTCAAATTCGAAGTATATAAT	1854
RESULT 12			
LOCUS	112242	2409 bp	DNA linear PAT 26-JUL-1995
DEFINITION	Sequence 18 from patent US 5422090.		
ACCESSION	112242		
VERSION	112242.1	GI:910265	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1. (bases 1 to 2409) Stephens,R.W., Golder,J.P., Antalio,T.M., Barnes,T.M., Clark,M.A., Devine,P.L., Goss,N.H. and Lehrbach,P.R.		
TITLE	Human PAI-2		
JOURNAL	Patent: US 5422090-A 18 06-JUN-1995;		
FEATURES	Location/Qualifiers		
source	1. 2409 /organism="unknown" /mol_type="unassigned DNA"		
ORIGIN			
Query Match	96.8%;	Score 1847.6;	DB 6; Length 2409;
Best Local Similarity	99.8%;	Pred. No. 0;	
Matches 1850;	Conservative 0;	Mismatches 4;	Indels 0; Gaps 0;
2y	15	GTCCAGACAGCAACTCAGAGAATACCCAGAGAACAACAGATTGAAACAATGGAGGATCTT	74
Db	1	GTCCAGACAGCAACTCAGAGAATACCCAGAGAACAACAGATTGAAACAATGGAGGATCTT	60
2y	75	TGTGTGGCAACACACACTTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCC	134
Db	61	TGTGTGGCAACACACACTTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCC	120
2y	135	ACCAGAAACCTTTCCTCCCGATGAGCATCTCGTCCACCATGGCCATGGTCTACATG	194
Db	121	ACCAGAAACCTTTCCTTCCTCCCGATGAGCATCTCGTCCACCATGGCCATGGTCTACATG	180
2y	195	GGCTCCAGGGGCGAGCACCGAAGACAGATGGCCAAAGTGCTTCAGTTTAATGAAGTGGGA	254
Db	181	GGCTCCAGGGGCGAGCACCGAAGACAGATGGCCAAAGTGCTTCAGTTTAATGAAGTGGGA	240
2y	255	GCCAAATCCAGTTACCCCGATGACATCCAGAGAACTTTTACCAGCTCTGGTTCATGCAGCAG	314
Db	241	GCCAAATCCAGTTACCCCGATGACATCCAGAGAACTTTACCAGCTCTGGTTCATGCAGCAG	300
2y	315	ATCCAGAAAGGGTAGTTATCTCGATGGCATTTTGGAGGCAAGCTGCAGATAAATCCAT	374
Db	301	ATCCAGAAAGGGTAGTTATCTCGATGGCATTTTGGAGGCAAGCTGCAGATAAATCCAT	360
2y	375	TCATCCCTTCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTATTTACTGGAA	434


```
Y 1455 ACAGAAATAATTAGACAAATTCCTATTATAAATGACAAACCTATTATAATCAATTTGGTCTT 1514
b 1441 ACAGAAATAATTAGACAAATTCCTATTATAAATGACAAACCTATTATAATCAATTTGGTCTT 1500
Y 1515 CTAAATGGGATCATGGCCATTTAGATTTTCCTACTATCAGTTTATTTTATTAACATTA 1574
b 1501 CTAAATGGGATCATGGCCATTTAGATTTTCCTACTATCAGTTTATTTTATTAACATTA 1560
Y 1575 ACTTTTACTTTGTTTATTATTATTATTATATATATATATATATATATATATATATATAT 1634
b 1561 ACTTTTACTTTGTTTATTATTATTATTATATATATATATATATATATATATATATATAT 1620
Y 1635 CCTATTTAATGACTAATAAGATTATAGAGCAGATGATCTGTTAAATTTCTATCTAAT 1694
b 1621 CCTATTTAATGACTAATAAGATTATAGAGCAGATGATCTGTTAAATTTCTATCTAAT 1680
Y 1695 AATGCTTTAAATTTGTTCTCATATATATATATATATATATATATATATATATATATATAT 1754
b 1681 AATGCTTTAAATTTGTTCTCATATATATATATATATATATATATATATATATATATATAT 1740
Y 1755 AATAAATATCTGGAAGAAACATTAACCAATAGGCAAAATATATATATATATATATATATAT 1814
b 1741 AATAAATATCTGGAAGAAACATTAACCAATAGGCAAAATATATATATATATATATATATAT 1800
Y 1815 AATACATAACACATATATATCTGCTATCTTATCTTATATCTAATTCAGATATATAAT 1868
b 1801 AATACATAACACATATATATCTGCTATCTTATCTTATATCTAATTCAGATATATAAT 1854

:RESULT 14
LOCUS A10503 1675 bp RNA linear PAT 30-AUG-1993
DEFINITION Artificial sequence for PAI-2.
ACCESSION A10503
VERSION A10503.1 GI:413574
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
FEATURES
     source
       1..1675
         /organism="synthetic construct"
         /mol_type="unassigned RNA"
         /db_xref="taxon:32630"
     CDS
       1..1086
         /note="Protein sequence is in conflict with the conceptual
         translation"
         /codon_start=1
         /transl_table=1
         /product="PAI-2"
         /protein_id="CAA00911.1"
         /db_xref="GI:4526844"
         /db_xref="EMBL:CAA00911"
         /translation="AKVLFQNEVGANAVTPTMPENTSCGFNQOIKGSPDAILQAQ
         AADKHSFRSLSSAINASTNGYLLESVNKLFGEKSASFREYIRLCQKYISEPQAV
         DFLECAEARKINSWVQTQTKIPNLLPEGSVDGTRMVLNVAIFYFKWKMTPEK
         KLNGLYPRVNSAORTPVQMYLRKLANIGYIEDLKAQILELPYAGDVSNFLLPDEI
         ADVSTGLELLSESI TYDKLNKWSKDKNAEDEVVYIPOKLEEHVELRSILRSMGE
         DAFVKGRANFSGMSERNDLFLSEVFQHMVDVNEEGTEAAAGTGGVMTGRTGHGQPF
         VADHPFLFLIMHKIINCILPFRFSP"

Query Match 87.3%; Score 1665.4; DB 6; Length 1675;
Best Local Similarity 99.6%; Pred. No.:0;
Matches 1669; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Y 225 GCCAAGGTGCTTCAGTTTAATGAAGTGGAGCAATGCAAGTACCCCACTGATCCAGAG 284
b 1 GCCAAGGTGCTTCAGTTTAATGAAGTGGAGCAATGCAAGTACCCCACTGATCCAGAG 60
```

```
QY 285 AACTTTACCAGCTCTGGGTTTCATGCAGCAGATCCAGAGGTAGTTTATCCTGTATGCGATT 344
DB 61 AACTTTACCAGCTCTGGGTTTCATGCAGCAGATCCAGAGGTAGTTTATCCTGTATGCGATT 120
QY 345 TTGCAGCAGCAAGCTGCAGATATAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATC 404
DB 121 TTGCAGCAGCAAGCTGCAGATATAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATC 180
QY 405 AATGCATCCACAGGGAATTTATTACTGAAAGTGTCATTAAGCTGTTTGGTGAAGTCT 464
DB 181 AATGCATCCACAGGGAATTTATTACTGAAAGTGTCATTAAGCTGTTTGGTGAAGTCT 240
QY 465 GCGAGCTTCGCGGGAAGATATATTTCGACTCTCTCAGAAATATTACTCCTCAGAAACCCCGAG 524
DB 241 GCGAGCTTCGCGGGAAGATATATTTCGACTCTCTCAGAAATATTACTCCTCAGAAACCCCGAG 300
QY 525 GCAGTAGACTTCCTAGATGTGCAGAAAGCTAGAAAAGATTAATTCCTGGGTCAAG 584
DB 301 GCAGTAGACTTCCTAGATGTGCAGAAAGCTAGAAAAGATTAATTCCTGGGTCAAG 360
QY 585 ACTCAAAACCAAGGCAAAATCCCAAACCTGTTTACCTGAAAGGTTCTGTAGATGGGATACC 644
DB 361 ACTCAAAACCAAGGCAAAATCCCAAACCTGTTTACCTGAAAGGTTCTGTAGATGGGATACC 420
QY 645 AGGATGCTCTGCTGATGCTGCTACTTCAAGGAAGTGAAGAACTCCATTTGAGAAG 704
DB 421 AGGATGCTCTGCTGATGCTGCTACTTCAAGGAAGTGAAGAACTCCATTTGAGAAG 480
QY 705 AAATAAATGGGCTTTATTCCTTTCCGCTGTAATCTGGCTCAGCGCACACCTGTACAGATG 764
DB 481 AAATAAATGGGCTTTATTCCTTTCCGCTGTAATCTGGCTCAGCGCACACCTGTACAGATG 540
QY 765 ATGTACTTGGCTGAAAAGCTAAACATTTGGATACATAGAAGACCTTAAAGGCTCAGATTCTA 824
DB 541 ATGTACTTGGCTGAAAAGCTAAACATTTGGATACATAGAAGACCTTAAAGGCTCAGATTCTA 600
QY 825 GAATCCCATATGCTGGAGATGTAGCATGTTCTGCTCCAGATGAATTTGCCGAT 884
DB 601 GAATCCCATATGCTGGAGATGTAGCATGTTCTGCTCCAGATGAATTTGCCGAT 660
QY 885 GTGTCCACTGGCTTTGGAGCTGCTGGAAGTGAATAAATCACTATGACAAATCAACAAGTGG 944
DB 661 GTGTCCACTGGCTTTGGAGCTGCTGGAAGTGAATAAATCACTATGACAAATCAACAAGTGG 720
QY 945 ACCAGCAAGACAAAATGGCTGAAGATGAAGTTGAGGTATACATACCCAGTTTCAAAATTA 1004
DB 721 ACCAGCAAGACAAAATGGCTGAAGATGAAGTTGAGGTATACATACCCAGTTTCAAAATTA 780
QY 1005 GAAGAGCATTTAAGAACTCAGATCCATTTCTGAAAAGCATGGGCATGGAGGACGCTTCAAC 1064
DB 781 GAAGAGCATTTAAGAACTCAGATCCATTTCTGAAAAGCATGGGCATGGAGGACGCTTCAAC 840
QY 1065 AAGGACCGGGCCAAATTTCTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTTCTGAAGTG 1124
DB 841 AAGGACCGGGCCAAATTTCTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTTCTGAAGTG 900
QY 1125 TTCCACCAAGCATGCTGGATGTGAATGAGAGGGGCACTGAAGCAGCCGCTGGCACAGGA 1184
DB 901 TTCCACCAAGCATGCTGGATGTGAATGAGAGGGGCACTGAAGCAGCCGCTGGCACAGGA 960
QY 1185 GGTGTTATGACAGGGAGAACTCGACATGGAGGCCCAAGTTTGTGGCAGATCATCCGTTT 1244
DB 961 GGTGTTATGACAGGGAGAACTCGACATGGAGGCCCAAGTTTGTGGCAGATCATCCGTTT 1020
QY 1245 CTTTTCCTTATTATGCATAAGATAACCAAGTGCATTTTATTTTTCGGCAGATTTTGTCTCA 1304
DB 1021 CTTTTCCTTATTATGCATAAGATAACCAAGTGCATTTTATTTTTCGGCAGATTTTTCCTCA 1080
QY 1305 CCCTAAAACCTAAGCGCTGCTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAG 1364
DB 1081 CCCTAAAACCTAAGCGCTGCTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCCTCAG 1140
```

1365 AATGCTATTTCAATGCGCAAAATTTAGAGATGTTTTCTACATATTTCTGCTCTCTG 1424
 1141 AATGCTATTTCAATGCGCAAAATTTAGAGATGTTTTCTACATATTTCTGCTCTCTG 1200
 1425 AACAACTTCTGCTACCCCACTAAATAAACAACAGAAATAATAGACAAATGCTCTATATA 1484
 1201 AACAACTTCTGCTACCCCACTAAATAAACAACAGAAATAATAGACAAATGCTCTATATA 1260
 1485 ACATGACAAACCCCTATTAATCAATTTGGTCTCTTAAATAGGATCATGCCCATTTAGATTTT 1544
 1261 ACATGACAAACCCCTATTAATCAATTTGGTCTCTTAAATAGGATCATGCCCATTTAGATTTT 1320
 1545 CTTTACTATGAGTTATTTTATTAACATTAATCTTTTACTTTGTTATTTATTTATTTATAT 1604
 1321 CTTTACTATGAGTTATTTTATTAACATTAATCTTTTACTTTGTTATTTATTTATTTATAT 1380
 1605 AATGCTGAGTTTTTAAATATTTGCTCACTGCTCTTAAATAGTACGTAATAAAGTTATAGA 1664
 1381 AATGCTGAGTTTTTAAATATTTGCTCACTGCTCTTAAATAGTACGTAATAAAGTTATAGA 1440
 1665 AGCAGATGATCTGTTATTTCTATCTATTAATAATGCTTTAATTTGTTCTCATTAATGAAGA 1724
 1441 AGCAGATGATCTGTTATTTCTATCTATTAATAATGCTTTAATTTGTTCTCATTAATGAAGA 1500
 1725 ATAAGTAGGTATCCCTCCATGCTCTTCTGTAATAAATATCTGGAATAAATCAATTAACAAT 1784
 1501 ATAAGTAGGTATCCCTCCATGCTCTTCTGTAATAAATATCTGGAATAAATCAATTAACAAT 1560
 1785 AGCAATATATGTTATGTTCTGATTTCTAGAAATACATACATATATATGCTGATCT 1844
 1561 AGCAATATATGTTATGTTCTGATTTCTAGAAATACATACATATATATGCTGATCT 1620
 1845 TATATTCATTTGCAAGTATATTAATAAATCACTGCTTCCAAACAACAAATAA 1899
 1621 TATATTCATTTGCAAGTATATTAATAAATCACTGCTTCCAAACAACAAATAA 1675

RESULT 15
 A21238
 LOCUS H.sapiens BTA 1445 mRNA for Pai-2.
 ACCESSION A21238
 VERSION A21238.1 GI:641356
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1610)

VARIANTS OF PAI-2
 Patent: WO 9109124-A 1 27-JUN-1991;
 Location/Qualifiers
 1..1610
 /organism="Homo sapiens"
 /mol_type="unassigned RNA"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18q21-q23"
 /clone="BTA 1445"
 /cell_line="U937"
 /cell_type="monocyte"
 49..1295
 /codon_start=1
 /product="plasminogen activator inhibitor type 2 protein"
 /protein_id="CAA01535.1"
 /db_xref="GI:641357"
 /db_xref="GOA:P05120"
 /db_xref="SWISS-PROT:P05120"
 /translation="MEDICVANTIFALNFKHLAKASPTQNLFLSPWSISSTMAWVY
 GRSRDEQWAKVLOFNEVGNAAVTMTPTNFTSCGFMQIQKSYDPDAILQAQADK
 IHSRPSLSAINASTGNVLLESVNKLFGEKSAFREYIRLCKYYSSSPQAVDFLE
 CAEARKKINSWKVTKGKI PNLLPGSDVDGTRMVLVNAVYFKGKWKTPFEKLLNG

CDS

LYPRVNSAORTPVMYLRKLNIGVIEDLKAQILELPYAGDVSMFLLLPDIADVS
 TGLELSEBITYDKLNKWTSDKWAEDVEVYIPOFLEHVELRSILRSMGMEAFN
 KGRANSGMSERNDLFLSEVPHOAMVDNEBTEAAAGTGGVMTGRTGGHGFQFVADH
 PFLFLNKHKITNCLLFFGRFSSP"

Query Match 84.0%; Score 1602; DB 6; Length 1610;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1605; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

ORIGIN

QY 15 GTGAGACAGCAACTCAGAGATAAACCAGAGAACCAACAGATTGAAACAATGGAGGATCTT 74
 DB 1 GTCAGACAGCAACTCAGAGATAAACCAGAGAACCAACAGATTGAAACAATGGAGGATCTT 60
 QY 75 TGTGTGCAACACACACTCTTTGCCCTCAATTTATTTCAAGCATCTCGCAAGACGACCC 134
 DB 61 TGTGTGCAACACACACTCTTTGCCCTCAATTTATTTCAAGCATCTCGCAAGACGACCC 120
 QY 135 ACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGCTCTACATG 194
 DB 121 ACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGCTCTACATG 180
 QY 195 GGCTCCAGGGGCGAGCAGGAGACAGATGCGCAAGGTGCTTCAGTTTAAATGAAGTGGGA 254
 DB 181 GGCTCCAGGGGCGAGCAGGAGACAGATGCGCAAGGTGCTTCAGTTTAAATGAAGTGGGA 240
 QY 255 GCCAATGAGTTACCCCATCACTCCAGAGAACTTTACAGCTGGGGTTCATGCGAGCAG 314
 DB 241 GCCAATGAGTTACCCCATCACTCCAGAGAACTTTTACAGCTGGGGTTCATGCGAGCAG 300
 QY 315 ATCCAGAGGGGTAGTTATCTCTGATCGGATTTTGCAAGGCAACAGCTGCAGATAAAATCCAT 374
 DB 301 ATCCAGAGGGGTAGTTATCTCTGATCGGATTTTGCAAGGCAACAGCTGCAGATAAAATCCAT 360
 QY 375 TCATCTTCCCTCTCTCAGCTCTCAATCAATGATGCATCCAGGGAAATTTTACTTGGAA 434
 DB 361 TCATCTTCCCTCTCTCAGCTCTCAATCAATGATGCATCCAGGGAAATTTTACTTGGAA 420
 QY 435 AGTGTCATAAGCTGTTTGTGAGAAAGTCTGCGAGCTTCCGGGAAGAATATATTTCGACTC 494
 DB 421 AGTGTCATAAGCTGTTTGTGAGAAAGTCTGCGAGCTTCCGGGAAGAATATATTTCGACTC 480
 QY 495 TGTGAGAAATATTTACTCTCTCAGAACCCAGGAGTAGACTTCTAGAAATGTCGAGAGAA 554
 DB 481 TGTGAGAAATATTTACTCTCTCAGAACCCAGGAGTAGACTTCTAGAAATGTCGAGAGAA 540
 QY 555 CCTAGAAAAAGATTAAATTTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTG 614
 DB 541 CCTAGAAAAAGATTAAATTTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTG 600
 QY 615 TTACTGGAAGGTCTCTGATAGGGGATACAGGATGCTCTGCTGATGCTGCTACTTTC 674
 DB 601 TTACTGGAAGGTCTCTGATAGGGGATACAGGATGCTCTGCTGATGCTGCTACTTTC 660
 QY 675 AAAGGAAAGTGGAAAACTCCATTTGAGAGAAACCTAAATGGGCTTTATCTCTTCGGTGA 734
 DB 661 AAAGGAAAGTGGAAAACTCCATTTGAGAGAAACCTAAATGGGCTTTATCTCTTCGGTGA 720
 QY 735 AACTCGGCTCAGCGCACACCTCTACAGATGATGATCTTGGCTGAAAAAGCTTAAACATTGGA 794
 DB 721 AACTCGGCTCAGCGCACACCTCTACAGATGATGATCTTGGCTGAAAAAGCTTAAACATTGGA 780
 QY 795 TACATAGAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGAGATGTTAGCATG 854
 DB 781 TACATAGAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGAGATGTTAGCATG 840
 QY 855 TTCTTTGTTGCTTCCAGATGAAATTTGCCGATGCTGCCACTGGCTTGGAGCTGCTGAAAGT 914
 DB 841 TTCTTTGTTGCTTCCAGATGAAATTTGCCGATGCTGCCACTGGCTTGGAGCTGCTGAAAGT 900
 QY 915 GAAATAAACCCTATGACAAAATCAACAACTCAACAACTGACCAAGCAACAAATGGCTGAGATGAA 974
 DB 901 GAAATAAACCCTATGACAAAATCAACAACTCAACAACTGACCAAGCAACAAATGGCTGAGATGAA 960

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

DM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 08:01:57 ; Search time 464.37 Seconds

(without alignments)
17454.974 Million cell updates/sec

Title: US-10-084-817-27

Perfect score: 1908

Sequence: 1 gagagcattgcccgtcaga.....aacaaaaataaaaaaagg 1908

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq 29Jan04: *
1: Geneseq1980s: *
2: Geneseq1990s: *
3: Geneseq2000a: *
4: Geneseq2001a: *
5: Geneseq2001b: *
6: Geneseq2002a: *
7: Geneseq2003a: *
8: Geneseq2003b: *
9: Geneseq2003c: *
10: Geneseq2004a: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1897	99.4	1897	ADA24501	ADA24501 Human CDN
2	1896	99.4	1935	AAH33195	AAH33195 Human col
3	1886	98.8	1920	AAH81019	AAH81019 Sequence
4	1877.2	98.4	1910	AAQ12590	AAQ12590 Placental
5	1872.6	98.1	1900	ADD18727	ADD18727 Human dis
6	1846	96.8	2409	AAH70474	AAH70474 Sequence
7	1750	91.7	2424	AAH91177	AAH91177 Sequence
8	1667	87.4	1675	AAH80060	AAH80060 Sequence
9	1405.6	73.7	1412	AAQ10849	AAQ10849 Encodes P
10	1402.6	73.5	1512	AAQ11819	AAQ11819 Plasminog
11	1341	70.3	1482	AAQ11820	AAQ11820 Plasminog
12	1249.4	65.5	1328	AAQ11128	AAQ11128 PreA-plas
13	1242.8	65.1	1340	AAQ11129	AAQ11129 PreB-PAI-
14	1242	65.1	1290	AAQ21119	AAQ21119 PAI-2 in
15	1014	53.1	1024	ABZ83797	ABZ83797 Toxicolog
16	984.8	51.6	1199	ABLS7801	ABLS7801 Oesophagu
17	567.8	29.8	601	AAH8579	AAH8579 Human chr
18	560	29.4	566	AAH56554	AAH56554 Human CDN
19	525.2	27.5	661	AAH72490	AAH72490 Single nu
20	525.2	27.5	661	AAH72484	AAH72484 Single nu
21	525.2	27.5	661	AAH72487	AAH72487 Single nu
22	524	27.5	661	AAH72451	AAH72451 Single nu
23	523.6	27.4	661	AAH72454	AAH72454 Single nu

24	523.6	27.4	661	3	AAH72496	AAH72496 Single nu
25	498.8	26.1	503	4	AAH56442	AAH56442 Human CDN
26	475.6	24.9	482	4	AAH35183	AAH35183 Probe #38
27	475.6	24.9	482	4	AAH29227	AAH29227 Human liv
28	475.6	24.9	482	4	AAH28847	AAH28847 Human liv
29	475.6	24.9	482	4	AAH28847	AAH28847 Human liv
30	315.4	16.5	317	4	AAH48342	AAH48342 Probe #17
31	315.4	16.5	317	4	AAH48342	AAH48342 Human bon
32	315.4	16.5	317	4	AAH48342	AAH48342 Human bon
33	307.4	16.1	316	2	AAH88576	AAH88576 Human chr
34	274.4	14.4	314	4	AAH56562	AAH56562 Human CDN
35	274.4	14.4	314	4	AAH56562	AAH56562 Human CDN
36	272.4	14.3	1278	8	AAH57843	AAH57843 Human CDN
37	272.4	14.3	1278	8	AAH57843	AAH57843 Human CDN
38	271	14.2	1200	7	AAH49851	AAH49851 CDNA enco
39	269	14.1	1185	6	AAH01980	AAH01980 Human ser
40	268.4	14.1	1316	2	AAH14255	AAH14255 Human NOV
41	268.4	14.1	1316	6	AAH69984	AAH69984 Pancreas
42	268.4	14.1	1316	6	AAH84121	AAH84121 Human CDN
43	268.4	14.1	1430	7	AAH56112	AAH56112 Horse sig
44	268.4	14.1	1458	6	AAH35236	AAH35236 Human gen
45	268.4	14.1	1550	9	AAH49056	AAH49056 Human NOV

ALIGNMENTS

RESULT 1
ADA24501
ID ADA24501 standard; CDNA; 1897 BP.
XX
AC ADA24501;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human CDNA differential expressed in adipose tissue, INCYTE063646CBL.
XX
KW ss; differential expression; adipose tissue; cytostatic; hypotensive;
KW antiarteriosclerotic; antidiabetic; anorectic; gene therapy;
KW peroxisome proliferator-activated receptor gamma; PPARGgamma;
KW diabetes mellitus; obesity; hypertension; atherosclerosis; breast cancer;
KW prostate cancer; colon cancer; polycystic ovarian syndrome.
XX
OS Homo sapiens.
XX
FN US2003096272-A1.
XX
PD 22-MAY-2003.
XX
PF 29-JUL-2002; 2002US-00208408.
XX
PR 30-JUL-2001; 2001US-0308868P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Schebye XM;
XX
WPI 2003-606416/57.
XX
PT New combination comprising several CDNAs, useful for preparing a
PT composition for diagnosing or treating diabetes mellitus, obesity,
PT hypertension, atherosclerosis, or cancer of the breast, prostate or
PT colon.
XX
PS Claim 1; Page 39-40; 84pp; English.
XX
CC Then invention relates to a new combination comprising 55 CDNAs (ADA24485
CC -ADA24539) or their complements that are differentially regulated in an
CC adipose sample. Also included are detecting differential expression of
CC one or more CDNAs in a sample containing nucleic acids, screening several
CC molecules or compounds to identify a ligand that specifically binds a
CC cDNA, a vector comprising the cDNA, a host cell containing the vector,
CC producing a protein, screening several molecules or compounds, producing

CC an antibody and the isolated antibody. The cDNAs comprise sequences which
 CC are upregulated or downregulated in response to peroxisome proliferator-
 CC activated receptor gamma (PPARGgamma) agonist. The combination comprising
 CC several cDNAs is useful for preparing a composition for diagnosing or
 CC treating diabetes mellitus, obesity, hypertension, atherosclerosis,
 CC cancer of the breast, prostate or colon, or polycystic ovarian syndrome.
 XX
 SQ Sequence 1897 BP; 595 A; 387 C; 382 G; 533 T; 0 U; 0 Other;

Query Match 99.4%; Score 1897; DB 8; Length 1897;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1897; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAGGAGCATGCGCGTCAGACAGCAACTCAGAGATTAACACAGAGACCAACAGATTGAAA	60
DB	1	GAGGAGCATGCGCGTCAGACAGCAACTCAGAGATTAACACAGAGACCAACAGATTGAAA	60
QY	61	CAATGAGGATCTTTGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGG	120
DB	61	CAATGAGGATCTTTGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGG	120
QY	121	CAAAAGCAAGCCCAACAGCAACTCTCTCTCCCATGAGCATCTCGTCCACCATGG	180
DB	121	CAAAAGCAAGCCCAACAGCAACTCTCTCTCCCATGAGCATCTCGTCCACCATGG	180
QY	181	CCATGCTCTACATGGGCTCCAGGGGCGACACCGAAGACAGATGCCCAAGGTGCTTCAGT	240
DB	181	CCATGCTCTACATGGGCTCCAGGGGCGACACCGAAGACAGATGCCCAAGGTGCTTCAGT	240
QY	241	TTAATGAAGTGGGAGCCAAATGCACTTACCCCATGACTCCAGAGAACTTTACCAAGCTGTG	300
DB	241	TTAATGAAGTGGGAGCCAAATGCACTTACCCCATGACTCCAGAGAACTTTACCAAGCTGTG	300
QY	301	GGTTTCATGACAGACATCCAGAGGGTATTTATCTGTATGCGATTTTGAGGCACAAAGCTG	360
DB	301	GGTTTCATGACAGACATCCAGAGGGTATTTATCTGTATGCGATTTTGAGGCACAAAGCTG	360
QY	361	CAGATAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATCCATCCACAGGA	420
DB	361	CAGATAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATCCATCCACAGGA	420
QY	421	ATTATTTACTGGAAAGTGTCAATAAGCTGTTTGTGTGAGAAGTCTCGAGCTTCGCGGAAG	480
DB	421	ATTATTTACTGGAAAGTGTCAATAAGCTGTTTGTGTGAGAAGTCTCGAGCTTCGCGGAAG	480
QY	481	AATATATTCGACTGTGAGAAATATTTACTCTCGAAGCCCGAGGATCTTCCTAG	540
DB	481	AATATATTCGACTGTGAGAAATATTTACTCTCGAAGCCCGAGGATCTTCCTAG	540
QY	541	AATGTGAGAAAGTGTAGAAAGAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCA	600
DB	541	AATGTGAGAAAGTGTAGAAAGAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCA	600
QY	601	AAATCCAAACTTGTACCTGAAGTTCTGTAGATGGGATACCAAGGATGGTCTGGTGA	660
DB	601	AAATCCAAACTTGTACCTGAAGTTCTGTAGATGGGATACCAAGGATGGTCTGGTGA	660
QY	661	ATGCTGTCTACTTCAAAAGGAAAGTGGAAAACCTCCATTTTGAGAAGAACTTAAATGGGCTTT	720
DB	661	ATGCTGTCTACTTCAAAAGGAAAGTGGAAAACCTCCATTTTGAGAAGAACTTAAATGGGCTTT	720
QY	721	ATCCTTTCCGTTGATACTCGGCTCAGCGCACCTGTGACAGATGATGTACCTGGTGAAA	780
DB	721	ATCCTTTCCGTTGATACTCGGCTCAGCGCACCTGTGACAGATGATGTACCTGGTGAAA	780
QY	781	AGCTAAACATTTGGATACATAGAAGACCTTAAAGGCTCAGATTTCTAGAACTCCCATATGCTG	840
DB	781	AGCTAAACATTTGGATACATAGAAGACCTTAAAGGCTCAGATTTCTAGAACTCCCATATGCTG	840
QY	841	GAGATGTAGCATGTTCTTTGTGCTTCAGATGAAATGCCGATGTGCTCCACTGGCTTGG	900
DB	841	GAGATGTAGCATGTTCTTTGTGCTTCAGATGAAATGCCGATGTGCTCCACTGGCTTGG	900

RESULT 2
 AAH33195
 ID AAH33195 standard; cDNA; 1935 BP.
 XX

QY	901	AGCTGCTGAAAGTGAAATTAACCTATGACAAACTCAACAAAGTGACACGACCAAGCAAAA	960
DB	901	AGCTGCTGAAAGTGAAATTAACCTATGACAAACTCAACAAAGTGACACGACCAAGCAAAA	960
QY	961	TGGCTCAAGATGAAGTTGAGGTATACATACCCAGTTCACAAATTAGAAGAGCATTTATGAAC	1020
DB	961	TGGCTCAAGATGAAGTTGAGGTATACATACCCAGTTCACAAATTAGAAGAGCATTTATGAAC	1020
QY	1021	TGAGATCCATTTCTGAAAAAGCATGGGATGAGAGAGCGCTTCAACAAAGGATGGGCAATT	1080
DB	1021	TGAGATCCATTTCTGAAAAAGCATGGGATGAGAGAGCGCTTCAACAAAGGATGGGCAATT	1080
QY	1081	TCTCAGGGATGTCGAGAGAGGAATGACCTGTTTCTTTCTGAGTGTTCACCAAGCCATGG	1140
DB	1081	TCTCAGGGATGTCGAGAGAGGAATGACCTGTTTCTTTCTGAGTGTTCACCAAGCCATGG	1140
QY	1141	TGATGTGAATGAGAGGGGCACTGAAGACGCGCTGACAGAGGTTTATGACAGGA	1200
DB	1141	TGATGTGAATGAGAGGGGCACTGAAGACGCGCTGACAGAGGTTTATGACAGGA	1200
QY	1201	GAACTGCATGAGAGGCCACAGTTCGTCAGATCATCCGTTCTTTCTTTCTTTTATATGC	1260
DB	1201	GAACTGCATGAGAGGCCACAGTTCGTCAGATCATCCGTTCTTTCTTTCTTTTATATGC	1260
QY	1261	ATAAGATAACCAAGTGCATTTTATTTTCGCGAGATTTTGTCTCACCCTAAACTTAAGCT	1320
DB	1261	ATAAGATAACCAAGTGCATTTTATTTTCGCGAGATTTTGTCTCACCCTAAACTTAAGCT	1320
QY	1321	GCTGCTCTTCGAAAAAGATTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTCAAT	1380
DB	1321	GCTGCTCTTCGAAAAAGATTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTCAAT	1380
QY	1381	TGCAAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGCAACAACTCTGCTACC	1440
DB	1381	TGCAAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGCAACAACTCTGCTACC	1440
QY	1441	CACATAAATAAACAACAGAAATTAATAGACAAATGTCTATTATATACATGACAAACCTATT	1500
DB	1441	CACATAAATAAACAACAGAAATTAATAGACAAATGTCTATTATATACATGACAAACCTATT	1500
QY	1501	AATCATTTGGTCTCTTAAATGGGATCATGCCCAATTTAGATTTTCTTACTATCAGTTTA	1560
DB	1501	AATCATTTGGTCTCTTAAATGGGATCATGCCCAATTTAGATTTTCTTACTATCAGTTTA	1560
QY	1561	TTTATTATAACATTAACCTTTTACTTTTATTATTATTATTATTATTAATGGTGAAGTTTAA	1620
DB	1561	TTTATTATAACATTAACCTTTTACTTTTATTATTATTATTATTATTAATGGTGAAGTTTAA	1620
QY	1621	ATTATTGCTCAGTGCCTTATTAATGAGTAAATTAAGTATAGAGCAGATGATCTGTTA	1680
DB	1621	ATTATTGCTCAGTGCCTTATTAATGAGTAAATTAAGTATAGAGCAGATGATCTGTTA	1680
QY	1681	ATTTCTCTATCTAATAAATGCTTTTAAATTTGTTCTCATTAATGAAGATAAGTAGTATCCCT	1740
DB	1681	ATTTCTCTATCTAATAAATGCTTTTAAATTTGTTCTCATTAATGAAGATAAGTAGTATCCCT	1740
QY	1741	CCATGCCCTTCTGTAATAAATATCTGAAAAACAATTAACAATAGGCAATATATGTTA	1800
DB	1741	CCATGCCCTTCTGTAATAAATATCTGAAAAACAATTAACAATAGGCAATATATGTTA	1800
QY	1801	TGTGCACTTTCTAGAAATACATACATATATGCTCTGTATCTTATATTCATTTGCAAG	1860
DB	1801	TGTGCACTTTCTAGAAATACATACATATATGCTCTGTATCTTATATTCATTTGCAAG	1860
QY	1861	TATATAAATAAATTAACCTGCTTCCAAACCAACAAAAA	1897
DB	1861	TATATAAATAAATTAACCTGCTTCCAAACCAACAAAAA	1897

AAH33195;
03-SEP-2001 (first entry)
Human colon cancer antigen encoding cDNA SEQ ID NO:251.
Human; colon cancer; colon cancer antigen; diagnosis; detection;
colorectal carcinoma; ss.
Homo sapiens.
WO200122920-A2.
05-APR-2001.
28-SEP-2000; 2000WO-US026524.
29-SEP-1999; 99US-0157137P.
03-NOV-1999; 99US-0163280P.
(HUMA-) HUMAN GENOME SCI INC.
Ruben SM, Barash SC, Birse CE, Rosen CA;
P-PSDB; AAG73764.
WPI; 2001-235357/24.
Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
useful for preventing, diagnosing and/or treating colorectal cancers.
Claim 1; Page 2385-2386; 9803pp; English.
AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
cancer-associated nucleic acid molecules (N) and proteins (P), where the
proteins are collectively known as colon cancer antigens. The colon
cancer antigens have cytostatic activity and can be used in gene therapy
and vaccine production. N and P may be used in the prevention, diagnosis
and treatment of diseases associated with inappropriate P expression. For
example, N and P may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of P by expressing inactive proteins or to
supplement the patient's own production of P. Additionally, N may be used
to produce the colon cancer-associated Ps, by inserting the nucleic acids
into a host cell and culturing the cell to express the proteins. N and P
can be used in the prevention, diagnosis and treatment of colorectal
carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent
sequences used in the exemplification of the present invention. N.B.
Pages 666 to 682 and page 7053 of the sequence listing were missing at
time of publication, meaning no sequences are present for SEQ ID NO:1027
to 1052, 7921 and 7922
Sequence 1935 BP; 611 A; 393 C; 392 G; 539 T; 0 U; 0 Other;
Query Match 99.4%; Score 1896; DB 4; Length 1935;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
3 GGAGCATTCGCGTCAGACGACACTCAGAGATAACCCAGACACACGATTTGAACA 62
15 GGAGCATTCGCGTCAGACGACACTCAGAGATAACCCAGACACACGATTTGAACA 74
63 ATGGAGGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGCA 122
75 ATGGAGGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGCA 134
123 AAAGCAAGCCCAACCCAGACACTCTCTCCCTCCCTGGAGCATCTCTCCACCATGGCC 182
135 AAAGCAAGCCCAACCCAGACACTCTCTCTCTCCCTGGAGCATCTCTCTCCACCATGGCC 194
183 ATGGTCTACATGGGCTCCAGGGGAGCAGCCGAGACACAGATGGCCAAAGGTCCTCAGTTT 242
195 ATGGTCTACATGGGCTCCAGGGGAGCAGCCGAGACACAGATGGCCAAAGGTCCTCAGTTT 254

QY 243 AATGAAGTGGAGCCCAATGCAGTTACCCCATGACTCCAGAGAACTTTTACCAGCTGTGGG 302
DB 255 AATGAAGTGGAGCCCAATGCAGTTACCCCATGACTCCAGAGAACTTTTACCAGCTGTGGG 314
QY 303 TTATGTCAGCAGATCCAGAAAGGTAGTTATCTCGATGCGATTTTTCAGGCAACAGCTGCA 362
DB 315 TTATGTCAGCAGATCCAGAAAGGTAGTTATCTCGATGCGATTTTTCAGGCAACAGCTGCA 374
QY 363 GATAAAATCCATTTCATCCCTTCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAAT 422
DB 375 GATAAAATCCATTTCATCCCTTCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAAT 434
QY 423 TATTACTGGAAGGTGTCATAAAGCTGTTTGGTGAGAAAGTCTCGAGCTTCGGGGAAGAA 482
DB 435 TATTACTGGAAGGTGTCATAAAGCTGTTTGGTGAGAAAGTCTCGAGCTTCGGGGAAGAA 494
QY 483 TATATTTCGACTCTGTCAGAAATATTACTCTCTCAGAACCCAGGAGTAGACTTCTCTAGAA 542
DB 495 TATATTTCGACTCTGTCAGAAATATTACTCTCTCAGAACCCAGGAGTAGACTTCTCTAGAA 554
QY 543 TGTGAGAAAGCTTAGAAAAAGATTAAATCTCGGTCAAGACTCAAAACCAAGGCAAA 602
DB 555 TGTGAGAAAGCTTAGAAAAAGATTAAATCTCGGTCAAGACTCAAAACCAAGGCAAA 614
QY 603 ATCCCAAACTTGTACTCAAGGTTCTGTAGATGGGATACCCAGGATGGTCTCTGTGTAAT 662
DB 615 ATCCCAAACTTGTACTCAAGGTTCTGTAGATGGGATACCCAGGATGGTCTCTGTGTAAT 674
QY 663 GCTGCTACTTCCAAAGGAAGTGGAAAACTCCATTTTGAAGAAAGAACTAAATTTGGGCTTTAT 722
DB 675 GCTGCTACTTCCAAAGGAAGTGGAAAACTCCATTTTGAAGAAAGAACTAAATTTGGGCTTTAT 734
QY 723 CTTTTCGGTGTAAACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTCCGTGAAAAG 782
DB 735 CTTTTCGGTGTAAACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTCCGTGAAAAG 794
QY 783 CTTAAACATTTGGATACATAGAAAGCCTTAAGGCTCAGATTTCTAGAACTCCCAATATCTGGA 842
DB 795 CTTAAACATTTGGATACATAGAAAGCCTTAAGGCTCAGATTTCTAGAACTCCCAATATCTGGA 854
QY 843 GATGTTAGCATGTTCTTGTGTTTCCAGATGAAATTTGCCGATGTTTCCACTGGCTTGGAG 902
DB 855 GATGTTAGCATGTTCTTGTGTTTCCAGATGAAATTTGCCGATGTTTCCACTGGCTTGGAG 914
QY 903 CTGCTGGAAGTGAATTAACCTATGACAACTCAAGTGGACACGAAAGACAAAATG 962
DB 915 CTGCTGGAAGTGAATTAACCTATGACAACTCAAGTGGACACGAAAGACAAAATG 974
QY 963 GCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAAGACATTTATGAATC 1022
DB 975 GCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAAGACATTTATGAATC 1034
QY 1023 AGATCCATTTCTGAAAAGCATGGGCATGGAGAGCGCTTCAACAGGAGCGGCCAATTC 1082
DB 1035 AGATCCATTTCTGAAAAGCATGGGCATGGAGAGCGCTTCAACAGGAGCGGCCAATTC 1094
QY 1083 TCAGGAGATCTCGAGAGGAATGACCTGTTTCTTTCTGAAGTGTTCACCAAGCCATGGTG 1142
DB 1095 TCAGGAGATCTCGAGAGGAATGACCTGTTTCTTTCTGAAGTGTTCACCAAGCCATGGTG 1154
QY 1143 GATGTGAATGAGAGGGCACTGAAGCAGCGCTGGGACAGAGGTGTTATCAGAGGAGA 1202
DB 1155 GATGTGAATGAGAGGGCACTGAAGCAGCGCTGGGACAGAGGTGTTATCAGAGGAGA 1214
QY 1203 ACTGGACATGGAGGGCCACAGATTTGTGGCAGATCATCCGTTTCTTTTCTTATTATGAT 1262
DB 1215 ACTGGACATGGAGGGCCACAGATTTGTGGCAGATCATCCCTTTCTTTCTTATTATGAT 1274
QY 1263 AAGATAACCAAGTGCATTTTATTTTTCGAGATTTTGTCTCACCCCTAAACCTAGCGTGC 1322
DB 1275 AAGATAACCAAGTGCATTTTATTTTTCGAGATTTTCTCTCACCCCTAAACCTAGCGTGC 1334
QY 1323 TGCTTCTGCAAAAGATTTTGTGATGAGTGTGTGCTCAGAAATGTGCTATTTCAAAATG 1382

```

1335 TGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAGAATGTGATTTCAAAATG 1394
1383 CCAAAATTTAGAGATGTTTCTACATATTTCTCTCTCTGTGAACAACTTCTGTACCCA 1442
1395 CCAAAATTTAGAGATGTTTCTACATATTTCTCTCTCTGTGAACAACTTCTGTACCCA 1454
1443 CTAATATATAAACAACAGAAATTAATAGCAATTTGCTATTATAACATGACCAACCTTATAA 1502
1455 CTAATATATAAACAACAGAAATTAATAGCAATTTGCTATTATAACATGACCAACCTTATAA 1514
1503 TCATTTGGTCTTAAATAGGATCATGCCATTTAGATTTTCTTACTATCATGATTTAT 1562
1515 TCATTTGGTCTTAAATAGGATCATGCCATTTAGATTTTCTTACTATCATGATTTAT 1574
1563 TTTATAACATTAACCTTTTACTTTTCTTATTTATTTATTTATTAATGCTGAGTTTTTAAAT 1622
1575 TTTATAACATTAACCTTTTACTTTTCTTATTTATTTATTTATTAATGCTGAGTTTTTAAAT 1634
1623 TATTGCTCAGTCTTATTAATGAGCTTAATAGATTAATAGATTTATAGAGCAGATGATCTGTTAT 1682
1635 TATTGCTCAGTCTTATTAATGAGCTTAATAGATTAATAGATTTATAGAGCAGATGATCTGTTAT 1694
1683 TTCTATCTAATAAATGCTTTTAAATTTCTTATAATGAAGAATAAGTAGGTATCCCTCC 1742
1695 TTCTATCTAATAAATGCTTTTAAATTTCTTATAATGAAGAATAAGTAGGTATCCCTCC 1754
1743 ATGCCCTTCTGATAAATATCTGGAACAAACATTAAGCAATATAGTATGTTATG 1802
1755 ATGCCCTTCTGATAAATATCTGGAACAAACATTAAGCAATATAGTATGTTATG 1814
1803 TGAATTTCTAGAAATACATAACATATATATGCTGTCTATCTTATTTCAATTCGAATG 1862
1815 TGAATTTCTAGAAATACATAACATATATATGCTGTCTATCTTATTTCAATTCGAATG 1874
1863 TATAATAAATAAATGCTTTTCAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1906
1875 TATAATAAATAAATGCTTTTCAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1918

RESULT 3
AA81019
ID AA81019 standard; cDNA; 1920 BP.
XX
AC AA81019;
XX
AC
XX
DT 25-MAR-2003 (revised)
DT 25-JAN-1991 (first entry)
XX
DE Sequence of plasmid pcD-1214 cDNA 1900 bp insert encoding human precursor
DE plasminogen activator inhibitor (PAI-2).
XX
KW Fibrinolytic therapy; diagnostic; epsilon-aminocaproic acid; probe; ss.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT sig_peptide 73..138
FT mat_peptide /*tag= a
FT polyA_signal 139..1320 /*tag= b
FT repeat_region 1456..1461 /*tag= c
FT repeat_region 1568..1577 /*tag= i
FT repeat_region /*note= "AT-rich repeat"
FT repeat_region 1597..1609 /*tag= j
FT repeat_region /*note= "AT-rich region"
FT repeat_region 1647..1653 /*tag= k
FT polyA_signal /*note= "AT-rich region"
FT 1661..1666

```

```

FT polyA_signal /*tag= d
FT 1702..1707 /*tag= e
FT polyA_signal 1764..1769 /*tag= f
FT polyA_signal 1875..1880 /*tag= g
FT polyA_signal 1879..1884 /*tag= h
XX
XX BP278696-A.
XX
XX 17-AUG-1988.
XX
XX 05-FEB-1988; 88EP-00300986.
XX
XX 06-FEB-1987; 87US-00011580.
XX
XX (NEW-) NEW ENGLAND MEDICAL CENT INC.
XX
XX (WELL-) WELLESLEY COLLEGE.
XX
XX (MASI) MIT.
XX
XX Webb AC, Auron PE;
XX
XX MPI; 1988-229534/33.
XX
XX P-PSDB; AAP80473.
XX
XX Human precursor and mature plasminogen activator inhibitor - obd. by
XX recombinant techniques from lipo-polysaccharide-stimulated human
XX monocytes.
XX
XX Disclosure; Fig 3a; 16pp; English.
XX
XX Human precursor plasminogen activator inhibitor (PAI-2) having specified
XX amino acid sequence and human mature PAI2 (23(S) to 415(P)) sequences are
XX claimed. Pure DNA coding for human precursor PAI-2 and human mature PAI-
XX 2, a transfer vector pcD1214 as available in E. coli HB101 (pcD1214)
XX (NRRL B-18161), transformed host cells (mammalian and yeast), and a
XX method of culturing the transformed cells to produce PAI-2 are claimed.
XX PAI-2 may have utility as a substitute for epsilon-aminocaproic acid
XX (epsilon-ACA). The PAI-2 gene can be used as a probe to locate the gene
XX coding for a PAI-2 or PAI-2-like activity. The PAI-2 can also be used to
XX produce antibodies used in diagnostic assays. (Updated on 25-MAR-2003 to
XX correct PA field.)
XX
XX Sequence 1920 BP; 612 A; 393 C; 380 G; 535 T; 0 U; 0 Other;
XX
XX Query Match 98.8%; Score 1886; DB 1; Length 1920;
XX Best Local Similarity 99.7%; Pred. No. 0;
XX Matches 1900; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 1 GAGGAGCATTGCCGTCAGACAGCAACTCAGAGATAACCCAGAGAACCAACCAAGATTGAAA 60
DB 11 GAGGAGCATTGCCGTCAGACAGCAACTCAGAGATAACCCAGAGAACCAACCAAGATTGAAA 70
QY 61 CAATGGAGGATCTTTGTGTGGCAACACACTTTTGGCCCTCAATTTATTCAGCATCTGG 120
DB 71 CAATGGAGGATCTTTGTGTGGCAACACACTTTTGGCCCTCAATTTATTCAGCATCTGG 130
QY 121 CAAAAGCAAGCCCAACCCAGAACCTCTTCTCTCCCAATGGAGCATCTCGTCCACCATGG 180
DB 131 CAAAAGCAAGCCCAACCCAGAACCTCTTCTCTCCCAATGGAGCATCTCGTCCACCATGG 190
QY 181 CCATGGTCTACATGGGCTCCAGGGGCGAGCCAGAGACAGATGGCAAGGTCCTTCAGT 240
DB 191 CCATGGTCTACATGGGCTCCAGGGGCGAGCCAGAGACAGATGGCAAGGTCCTTCAGT 250
QY 241 TTAATGAAGTGGAGCCCAATGCAGTTACCCCATCACTCCAGAGAACTTTACAGCTGTG 300
DB 251 TTAATGAAGTGGAGCCCAATGCAGTTACCCCATCACTCCAGAGAACTTTACAGCTGTG 310
QY 301 GGTTTCATGCAGATCCAGAGGGGTAGTTATCTCGATGCGATTTTTCAGGCAAGCTG 360

```

311 GGTTCATGCGACGATCAGAAAGGTAGTTATCCTGATCGGATTTTGCAGGCACAGCTG 370
361 CAGATAAAATCCATTCCTTCCTCCCTCTCTCAGCTCTGCAATCAATGATCAGAGGA 420
371 CAGATAAAATCCATTCCTTCCTCCCTCTCTCAGCTCTGCAATCAATGATCAGAGGA 430
421 ATATTTACTGGAAGTGTCAATAAGCTGTGTTGGTGAGAGTCTGCGAGCTTCCGGGAAG 480
431 ATATTTACTGGAAGTGTCAATAAGCTGTGTTGGTGAGAGTCTGCGAGCTTCCGGGAAG 490
481 AATATATTCGACTCTGTCAGAAATATTAATCTCTCAGAACCCAGGCGAGTAGACTTCCCTAG 540
491 AATATATTCGACTCTGTCAGAAATATTAATCTCTCAGAACCCAGGCGAGTAGACTTCCCTAG 550
541 AATCTGCAAGAAGAGCTAGAAAAGATTAATTCCTGGGTCAAGTCAAAACCAAGGCA 600
551 AATCTGCAAGAAGAGCTAGAAAAGATTAATTCCTGGGTCAAGTCAAAACCAAGGCA 610
601 AATATCCCAAACTTGTACTGAAGTTCGTAGATGGGATACCGAGTGGTCTGGTGA 660
611 AATATCCCAAACTTGTACTGAAGTTCGTAGATGGGATACCGAGTGGTCTGGTGA 670
661 ATGCTGTCTACTTCAAGAAAGTGAAGAACTCCATTTGAGAGAAACTAAATGGGCTTT 720
671 ATGCTGTCTACTTCAAGAAAGTGAAGAACTCCATTTGAGAGAAACTAAATGGGCTTT 730
721 ATCTTTTCCGTGTAACCTCGGCTCAGCGCACACCTGTACAGATGATGTACTTCCGTGAAA 780
731 ATCTTTTCCGTGTAACCTCGGCTCAGCGCACACCTGTACAGATGATGTACTTCCGTGAAA 790
781 AGCTAAACATTTGGATACATAGAGACCTAAAGCTCAGATCTAGAACCTCCATATGCTG 840
791 AGCTAAACATTTGGATACATAGAGACCTAAAGCTCAGATCTAGAACCTCCATATGCTG 850
841 GAGATGTTAGCATGTTCTTGTGTTTCCAGATGAAATTCGCGATGTTCACCTGGCTGG 900
851 GAGATGTTAGCATGTTCTTGTGTTTCCAGATGAAATTCGCGATGTTCACCTGGCTGG 910
901 AGCTGCTGGAAGTGAATAAATCTATGACAACTCAACAGTGCAGCAGCAAGACAAAA 960
911 AGCTGCTGGAAGTGAATAAATCTATGACAACTCAACAGTGCAGCAGCAAGACAAAA 970
961 TGGCTGAAGATGAAGTTAGGTATACATACCCAGTTCAAAATTAGAGAGCATTATGAAC 1020
971 TGGCTGAAGATGAAGTTAGGTATACATACCCAGTTCAAAATTAGAGAGCATTATGAAC 1030
1021 TCAGATCCATCTGAAAAGCATGGGCATGGAGACGCTTCAACAGGGGACGGGCCAATT 1080
1031 TCAGATCCATCTGAAAAGCATGGGCATGGAGACGCTTCAACAGGGGACGGGCCAATT 1090
1081 TCTCAGGATGTCGAGAGGAATGACCTGTTTCTTCTGAGTGTTCACCAAGCCATGG 1140
1091 TCTCAGGATGTCGAGAGGAATGACCTGTTTCTTCTGAGTGTTCACCAAGCCATGG 1150
1141 TGGATGTGAATGAGAGGGGACCTGAAAGCAGCCGCTGGCAGAGGAGTGTATGACAGGGA 1200
1151 TGGATGTGAATGAGAGGGGACCTGAAAGCAGCCGCTGGCAGAGGAGTGTATGACAGGGA 1210
1201 GAACTGGACATGGAGGCCACACAGTTGTGGCAGATCATCCGTTCTTTCTTATTATGC 1260
1211 GAACTGGACATGGAGGCCACACAGTTGTGGCAGATCATCCCTTTCTTATTATGC 1270
1261 ATAAGATACCAAGTGCATTTTATTTTTCGGCAGATTTTCTCACCCTAAACCTAAGCGT 1320
1271 ATAAGATACCAAGTGCATTTTATTTTTCGGCAGATTTTCTCACCCTAAACCTAAGCGT 1330
1321 GCTGCTCTGCAAAAGATTTTCTAGATGAGTGTGGCTCAGAAATTCCTATTTCAAAT 1380
1331 GCTGCTCTGCAAAAGATTTTCTAGATGAGTGTGGCTCAGAAATTCCTATTTCAAAT 1390
1381 TGCCAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAATTTCTGCTACC 1440
1391 TGCCAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAATTTCTGCTACC 1450

QY 1441 CACTAATAAACAACAGAAATAATTAGACAATTCCTATTATAACATGACAAACCCCTATT 1500
Db 1451 CACTAATAAACAACAGAAATAATTAGACAATTCCTATTATAACATGACAAACCCCTATT 1510
QY 1501 AATCATTTGGTCTTCTAAAATGGGATCATGCCATTTAGATTTTCTTACTACTACAGTTTA 1560
Db 1511 AATCATTTGGTCTTCTAAAATGGGATCATGCCATTTAGATTTTCTTACTACTACAGTTTA 1570
QY 1561 TTTTATAAACAATTAACCTTTTACTTTTGTATTATTATTATTATTATATATGTTGAGTTTTTAA 1620
Db 1571 TTTTATAAACAATTAACCTTTTACTTTTGTATTATTATTATTATTATATATGTTGAGTTTTTAA 1630
QY 1621 ATTATGCTCTACTGCTTATTTAAATGCTAGCTAATAAGTTTATAGAAGCAGATGATCTGTTA 1680
Db 1631 ATTATGCTCTACTGCTTATTTAAATGCTAGCTAATAAGTTTATAGAAGCAGATGATCTGTTA 1690
QY 1681 ATTTCTCTATCTAATAAATGCCCTTTAAATGCTCTAATAAATGAAGAAATAGTAGTATCCCT 1740
Db 1691 ATTTCTCTATCTAATAAATGCCCTTTAAATGCTCTAATAAATGAAGAAATAGTAGTATCCCT 1749
QY 1741 CCATGCCCTCTGTAATAAATATCTGGAATAAATCAATTAAGCAATAGGCAATATATGTTA 1800
Db 1750 CCATGCCCTCTGTAATAAATATCTGGAATAAATCAATTAAGCAATAGGCAATATATGTTA 1809
QY 1801 TGTGATTTCTAGAAATACATAACATATATATGCTGTATCTTATTTCAATTTGCAAG 1860
Db 1810 TGTGATTTCTAGAAATACATAACATATATATGCTGTATCTTATTTCAATTTGCAAG 1869
QY 1861 TATATAATAAATAAACCTGCTTCCAAACAACAAAAAATAAAAAA 1906
Db 1870 TATATAATAAATAAACCTGCTTCCAAACAACAAAAAATAAAAAA 1915

RESULT 4
AAQ12590
ID AAQ12590 standard; cDNA; 1910 BP.
XX
AC AAQ12590;
XX
DT 25-MAR-2003 (revised)
DT 26-SEP-1991 (first entry)
XX Placental PAI-2.
DE
XX Plasmidogen activator inhibitor; anticoagulant; ss.
XX Homo sapiens.
FH Key
CDS Location/Qualifiers
FT 56..1303
FT /*tag= a
FT polyA_signal 1439
FT /*tag= b
FT polyA_signal 1644
FT /*tag= c
FT polyA_signal 1685
FT /*tag= d
FT polyA_signal 1748
FT /*tag= e
FT polyA_signal 1859
FT /*tag= f
FT polyA_signal 1863
FT /*tag= g
FT polyA_site 1885..1910
FT /*tag= h
XX US028534-A.
PN
XX
XX 02-JUL-1991.
XX 15-SEP-1987; 87US-00097482.

PR 15-SEP-1987; 87US-00097482.
 CX (UNIW) UNIV WASHINGTON.
 PA (MONS) MONSANTO CO.
 KX
 PI Sadler JE, Wun TC;
 DR WPI, 1991-215146/29.
 DR P-FSDB; AAR13007.
 XX
 XX Deoxyribonucleic acid - encoding recombinant human plasminogen activator
 PT inhibitor, used for preventing conversion of plasminogen to plasmin.
 PT
 XX
 PS Claim 1; Fig 4; 23pp; English.
 XX
 XX The sequence was obtd. from the combined sequences of lambda PAI-75.1 and
 CC lambda PAI75.15, both isolated from a placental cDNA library. A
 CC recombinant vector contg. the DNA can be used to express PAI-2 which
 CC inhibits the conversion of plasminogen to plasmin and is used as an
 CC anticoagulant. (Updated on 25-MAR-2003 to correct PA field.)
 XX
 XX Sequence 1910 BP; 611 A; 382 C; 384 G; 533 T; 0 U; 0 Other;
 SQ
 Query Match 98.4%; Score 1877.2; DB 2; Length 1910;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1885; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 9 TTGCGCGTCAGACAGCACTCAGAGATACCGAGAACACAGATGGAACATGGAG 68
 DB 2 TTACCGCGTCAGACAGCACTCAGAGATACCGAGAACACAGATGGAACATGGAG 61
 QY 69 GATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGGCAAAAGCA 128
 DB 62 GATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGGCAAAAGCA 121
 QY 129 AGCCCAACCCAGAACCTCTTCTCTCCCGATGGAGCATCTGTCACCATGGCCATGGTC 188
 DB 122 AGCCCAACCCAGAACCTCTTCTCTCCCGATGGAGCATCTGTCACCATGGCCATGGTC 181
 QY 189 TACATGGGCTCCAGGGCAGCAGCAGAACAGATGCGCAAGGTGCTTCAGTTTAATGAA 248
 DB 182 TACATGGGCTCCAGGGCAGCAGCAGAACAGATGCGCAAGGTGCTTCAGTTTAATGAA 241
 QY 249 GTGGGAGCCAAATCAGTTACCCCATGACCTCCAGAGAACTTTTACAGCTGTGGGTTTCATG 308
 DB 242 GTGGGAGCCAAATCAGTTACCCCATGACCTCCAGAGAACTTTTACAGCTGTGGGTTTCATG 301
 QY 309 CAGCAGATCCAGAGGTAGTTATCTCTGATGCGATTTTGCAGGCACAGCTCCAGATBAA 368
 DB 302 CAGCAGATCCAGAGGTAGTTATCTCTGATGCGATTTTGCAGGCACAGCTCCAGATBAA 361
 QY 369 ATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCAGAGGATTAATTA 428
 DB 362 ATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCAGAGGATTAATTA 421
 QY 429 CTGGAAGTCTCAATAGCTTTTGTGAGAGCTGCGAGCTTCCGGAGAGAAATATTT 488
 DB 422 CTGGAAGTCTCAATAGCTTTTGTGAGAGCTGCGAGCTTCCGGAGAGAAATATTT 481
 QY 489 CGACTCTGTGAGAAATATTACTCTCGAAGCCCGAGGAGTAGACTTCTTAGAATGTGCA 548
 DB 482 CGACTCTGTGAGAAATATTACTCTCGAAGCCCGAGGAGTAGACTTCTTAGAATGTGCA 541
 QY 549 GAAGAGCTGAGAAAAGATTAAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCA 608
 DB 542 GAAGAGCTGAGAAAAGATTAAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCA 601
 QY 609 AACTTGTACCTGAAGGTTCTGTAGATGGGATACAGAGTGGTCTGGTGAATGCTGTC 668
 DB 602 AACTTGTACCTGAAGGTTCTGTAGATGGGATACAGAGTGGTCTGGTGAATGCTGTC 661
 QY 669 TACTTCAAGGAAAGTGGAAAACCTCCATTTGAGAGAAACTAAATGGGCTTTATCCTTTC 728

DB 562 TACTTCAAGGAAAGTGGAAAACCTCCATTTGAGAGAAACTAAATGGGCTTTATCTCTTC 721
 QY 729 CGTGTAAACTCGGCTCAGCGCACACTGTACAGATGATCTACTGCTGGAAGCTAAAC 788
 DB 722 CGTGTAAACTCGGCTCAGCGCACACTGTACAGATGATCTACTGCTGGAAGCTAAAC 781
 QY 789 ATTGGATACATAGAGACCTAAAGGCTCAGATTTCTAGAACTCCCATATGCTGGAGATGTT 848
 DB 782 ATTGGATACATAGAGACCTAAAGGCTCAGATTTCTAGAACTCCCATATGCTGGAGATGTT 841
 QY 849 AGCATGTTCTTTGCTCTCCAGATGAAATTTGCGATGTGTCCACTGGCTTGGAGCTGCTG 908
 DB 842 AGCATGTTCTTTGCTCTCCAGATGAAATTTGCGATGTGTCCACTGGCTTGGAGCTGCTG 901
 QY 909 GAAAGTGAATAACCTATGACAAAACCTCAACAAGTGCACGCAAGCAAAATGGCTGAA 968
 DB 902 GAAAGTGAATAACCTATGACAAAACCTCAACAAGTGCACGCAAGCAAAATGGCTGAA 961
 QY 969 GATGAAGTTGAGGTATACATACCCAGTTTCAAAATTAGAAGAGCATTTATGAATCAGATCC 1028
 DB 962 GATGAAGTTGAGGTATACATACCCAGTTTCAAAATTAGAAGACCATTTATGAATCAGATCC 1021
 QY 1029 ATTCTGAAAAGCATGGGATGAGGACGCTTCAACAAGGAGGGGCCAAATTTCTCAGG 1088
 DB 1022 ATTCTGAAAAGCATGGGATGAGGACGCTTCAACAAGGAGGGGCCAAATTTCTCAGG 1081
 QY 1089 ATGTCCGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATGTTGGATGTG 1148
 DB 1082 ATGTCCGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATGTTGGATGTG 1141
 QY 1149 AATGAGGAGGACCTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGAGAACTCGA 1208
 DB 1142 AATGAGGAGGACCTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGAGAACTCGA 1201
 QY 1209 CATGAGGCCCAACAGTTTGTGGCAGATCATCCGTTCTTTTCTTCTTATATGCAATAGATA 1268
 DB 1202 CATGAGGCCCAACAGTTTGTGGCAGATCATCCGTTCTTTTCTTCTTATGCAATAGATA 1261
 QY 1269 ACCAAGTGCATTTATTTTTCGSCAGATTTTGTCTCACCTAAAACTAAGCGTGTGCTTC 1328
 DB 1262 ACCAAGTGCATTTATTTTTCGSCAGATTTTGTCTCACCTAAAACTAAGCGTGTGCTTC 1321
 QY 1329 TCGAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATTTGCTATTTCAAAATGCGCAAAA 1388
 DB 1322 TCGAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATTTGCTATTTCAAAATGCGCAAAA 1381
 QY 1389 ATTAGAGATGTTTCTTACATATTTCTGCTCTTCTGAAACAACTTCTGCTACCCATAAT 1448
 DB 1382 ATTAGAGATGTTTCTTACATATTTCTGCTCTTCTGAAACAACTTCTGCTACCCATAAT 1441
 QY 1449 AAAACACAGAAATTAATAGCAATTTGCTATTTATACATGACAAACCTATTATCAATTT 1508
 DB 1442 AAAACACAGAAATTAATAGCAATTTGCTATTTATTAACATGACAAACCTATTATCAATTT 1501
 QY 1509 GGTCTTCTTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATTTTATA 1568
 DB 1502 GGTCTTCTTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATTTTATA 1561
 QY 1569 ACATTAACCTTTACTTTTGTATTATTTATTTATTAATGCTGAGTTTAAATTTATTTGC 1628
 DB 1562 ACATTAACCTTTACTTTTGTATTATTTATTTATTAATGCTGAGTTTAAATTTATTTGC 1621
 QY 1629 TCATGCGCTATTTAATGATAGCTAATAAGTTATAGAAGCAGATGATCTGTTAATTTCCCTA 1688
 DB 1622 TCATGCGCTATTTAATGATAGCTAATAAGTTATAGAAGCAGATGATCTGTTAATTTCCCTA 1681
 QY 1689 TCTAAATAAATGCGCTTTTAAATTTGTTCTCATATTAATGAAGAATAAGTAGGTATCCCTCCATGCC 1748
 DB 1682 TCTAAATAAATGCGCTTTTAAATTTGTTCTCATATTAATGAAGAATAAGTAGGTATCCCTCCATGCC 1741
 QY 1749 TTTCTGTAATTAATATCTGGAAAACCAATTAACAATAGGCAAAATATATGTTATGTGCATTT 1808
 DB 1742 TTTCTATAATAATATCTGGAAAACCAATTAACAATAGGCAAAATATATGTTATGTGCATTT 1801

1809 TCTAGAAATACATAACACATATATATCTCTCTATATATTTCAATTCGAAGTATATAAT 1868
 1802 TCTAGAAATACATAACACATATATATCTCTCTATATATTTCAATTCGAAGTATATAAT 1861
 1869 AAATAAAACCTGCTTCCAAACACCAAAAAATAAAAAA 1906
 1862 AAATAAAACCTGCTTCCAAACCAAGAAAAAATAAAAAA 1899

RESULT 5
 DD18727
 D ADD18727 standard; DNA; 1900 BP.
 X ADD18727;
 X AC
 X CX
 X JT 15-JAN-2004 (first entry)
 X CX
 X CX
 X CX Human disease related protein DNA sequence SeqID158.
 X CX human; disease state; cytostatic; antiinflammatory; ophthalmological;
 X CX antiarteriosclerotic; vulnerary; gene therapy; angiogenesis; apoptosis;
 X CX hypoxia-regulated condition; tumorigenesis; gliuconeogenesis;
 X CX inflammation; erythropoiesis; glycolysis; catecholamine synthesis;
 X CX glucose transportation; catecholamine synthesis; iron transport;
 X CX nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
 X CX retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
 X CX inflammatory condition; wound healing; gene; ds.
 X CX Homo sapiens.
 X CX WO2003018621-A2.
 X CX 06-MAR-2003.
 X CX 23-AUG-2002; 2002WO-GB003892.
 X CX 23-AUG-2001; 2001GB-00020538.
 X CX 05-OCT-2001; 2001GB-00024037.
 X CX (OXFO-) OXFORD BIOMEDICA UK LTD.
 X CX Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
 X CX WPI; 2003-290046/28.
 X CX P-PSDB; ADD18726.
 X CX New substantially purified polypeptide, useful for diagnosing or treating
 X CX a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
 X CX injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
 X CX wound healing.
 X CX Claim 27; SEQ ID NO 158; 424pp; English.
 X CX This invention relates to novel human genes and gene product which are
 X CX implicated in certain disease states. Compounds which modulate the
 X CX proteins of the invention may have cytostatic, antiinflammatory, the
 X CX ophthalmological, antiarteriosclerotic or vulnerary activities. The
 X CX sequences of the invention may be useful for gene therapy. The invention
 X CX may be useful for diagnosing or treating a hypoxia-regulated condition,
 X CX such as tumorigenesis, angiogenesis, apoptosis, inflammation,
 X CX erythropoiesis, or the biological response to hypoxia conditions
 X CX including processes such as glycolysis, gluconeogenesis, glucose
 X CX transportation, catecholamine synthesis, iron transport or nitric oxide
 X CX synthesis. The disease includes cancer, ischaemic conditions, reperfusion
 X CX injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
 X CX inflammatory conditions or wound healing. The present sequence is that of
 X CX a disease related protein encoding DNA sequence of the invention.
 X CX Sequence 1900 BP; 592 A; 393 C; 380 G; 535 T; 0 U; 0 Other;
 X CX Query Match 98.1%; Score 1872.6; DB 9; Length 1900;
 X CX Best Local Similarity 99.7%; Pred. No. 0;

		Matches	1886;	Conservative	0;	Mismatches	4;	Indels	1;	Gaps	1;
Qy	1	GAGGAGCAATTC	CGGTG	CAGACG	CACTC	CAGAGA	TAAC	CCAGAGA	CAAC	CCAGATTC	GA
Db	11	GAGGAGCAATTC	CGGTG	CAGACG	CACTC	CAGAGA	TAAC	CCAGAGA	CAAC	CCAGATTC	GA
Qy	61	CAATGGAGGAT	CTTTG	TGTG	CAAA	CACACT	CTTTG	CCCTCA	ATTTAT	TTCAGAGCAT	CG
Db	71	CAATGGAGGAT	CTTTG	TGTG	CAAA	CACACT	CTTTG	CCCTCA	ATTTAT	TTCAGAGCAT	CG
Qy	121	CAAAAGCAAG	CCCCA	CCAGAA	CCCTT	CTCTC	CTCC	CAATGG	AGCATCT	CTCCACCAT	GG
Db	131	CAAAAGCAAG	CCCCA	CCAGAA	CCCTT	CTCTC	CTCC	CAATGG	AGCATCT	CTCCACCAT	GG
Qy	181	CCATGGTCTA	CATGG	CTCC	AGGG	CAGCA	CCGAA	GACAGAT	GGCCAG	GTGCTT	CAGT
Db	191	CCATGGTCTA	CATGG	CTCC	AGGG	CAGCA	CCGAA	GACAGAT	GGCCAG	GTGCTT	CAGT
Qy	241	TTAATGAAG	TGGAG	CCCAAT	GCAGTT	ACCC	CCATG	ACTCC	AGAGAA	CTTTAC	CAGCTGTG
Db	251	TTAATGAAG	TGGAG	CCCAAT	GCAGTT	ACCC	CCATG	ACTCC	AGAGAA	CTTTAC	CAGCTGTG
Qy	301	GGTTTCATG	CAGCAG	ATCC	CAGAG	GGGTAG	TTAT	CTCGAT	GCATTT	TTCAGG	CAACAGCTG
Db	311	GGTTTCATG	CAGCAG	ATCC	CAGAG	GGGTAG	TTAT	CTCGAT	GCATTT	TTCAGG	CAACAGCTG
Qy	361	CAGATAAAT	CCATTC	ATCTT	CCGCT	CTCTC	AGCT	CTCTG	CAATCA	TGCAAT	GCATCCACAGGA
Db	371	CAGATAAAT	CCATTC	ATCTT	CCGCT	CTCTC	AGCT	CTCTG	CAATCA	TGCAAT	GCATCCACAGGA
Qy	421	ATTATTTAC	TGGAAG	GTG	CAATA	AGCTG	TTT	GGTGAG	AGTCT	GGAGCTT	CCGGGAAG
Db	431	ATTATTTAC	TGGAAG	GTG	CAATA	AGCTG	TTT	GGTGAG	AGTCT	GGAGCTT	CCGGGAAG
Qy	481	AATATATT	CGACT	CTGT	CAGAA	TATTA	CTCT	CCTCAG	AAACCC	CGGCGAG	TAGATTCCTAG
Db	491	AATATATT	CGACT	CTGT	CAGAA	TATTA	CTCT	CCTCAG	AAACCC	CGGCGAG	TAGATTCCTAG
Qy	541	AATGTGCA	GAGAAG	ACT	TAGAAA	AGATT	AAAT	TCTTGG	GTCAAG	ACTCAAA	CCAAAGGCA
Db	551	AATGTGCA	GAGAAG	ACT	TAGAAA	AGATT	AAAT	TCTTGG	GTCAAG	ACTCAAA	CCAAAGGCA
Qy	601	AAATCCCA	AACTT	GTG	TACT	CGAG	GGT	CTGT	TAGAT	GGGAT	TACAGGATGGTCTGGTGA
Db	611	AAATCCCA	AACTT	GTG	TACT	CGAG	GGT	CTGT	TAGAT	GGGAT	TACAGGATGGTCTGGTGA
Qy	661	ATGCTGTCT	ACTT	CAAG	AAAG	AGTGG	AAAC	CTT	TGAG	AAAGAA	CTAAATGGGCTTT
Db	671	ATGCTGTCT	ACTT	CAAG	AAAG	AGTGG	AAAC	CTT	TGAG	AAAGAA	CTAAATGGGCTTT
Qy	721	ATCCTTTCC	GGTAA	AACT	CGGCT	CAGC	GCAC	ACCT	GTG	TACAG	ATGATGTACTTGGGTGAAA
Db	731	ATCCTTTCC	GGTAA	AACT	CGGCT	CAGC	GCAC	ACCT	GTG	TACAG	ATGATGTACTTGGGTGAAA
Qy	781	AGCTAAAC	ATTGG	ATAC	TAGAA	GACCT	TAAG	AGCT	CAGAT	TCTAG	AACTCCCATATGCTG
Db	791	AGCTAAAC	ATTGG	ATAC	TAGAA	GACCT	TAAG	AGCT	CAGAT	TCTAG	AACTCCCATATGCTG
Qy	841	GAGATGTT	AGCAT	GTG	TTT	GTGCT	TCC	AGAT	GAAAT	TGCGAT	GTGTCTCCACTGGCTGG
Db	851	GAGATGTT	AGCAT	GTG	TTT	GTGCT	TCC	AGAT	GAAAT	TGCGAT	GTGTCTCCACTGGCTGG
Qy	901	AGCTCTG	GAAGT	GAATA	AACT	CTAT	GAC	AAAT	CAAC	AGTGG	ACCCAGGAAAGCAAAA
Db	911	AGCTCTG	GAAGT	GAATA	AACT	CTAT	GAC	AAAT	CAAC	AGTGG	ACCCAGGAAAGCAAAA
Qy	961	TGGCTG	CAGAT	GAAGT	TTGAG	GTAT	TAC	ATAC	ATCC	AGGTTCA	AAATTTAGAGAGCATTTATGAC
Db	971	TGGCTG	CAGAT	GAAGT	TTGAG	GTAT	TAC	ATAC	ATCC	AGGTTCA	AAATTTAGAGAGCATTTATGAC
Qy	1021	TCAGATCC	CAATTC	TGAA	AAAG	CAATGG	CGCAT	GGG	CGCTT	CAAC	AGGAGCGGGCAATT
Db	1031	TCAGATCC	CAATTC	TGAA	AAAG	CAATGG	CGCAT	GGG	CGCTT	CAAC	AGGAGCGGGCAATT

QY 1081 TCTCAGGATGTCGAGAGGAATGACCTGTTTCTTCTGAAGTGTTCACCAAGCCATGG 1140
 Db 1091 TCTCAGGATGTCGAGAGGAATGACCTGTTTCTTCTGAAGTGTTCACCAAGCCATGG 1150
 QY 1141 TGGATGTGAATGAGAGGGCACTGAAGCAGCGCTGGCAGAGGTTTATCAGAGGA 1200
 Db 1151 TGGATGTGAATGAGAGGGCACTGAAGCAGCGCTGGCAGAGGTTTATCAGAGGA 1210
 QY 1201 GAACCTGGACATGGAGGCCACACAGTTGTGGCAGATCATCGTTCCTTTCTTATATGC 1260
 Db 1211 GAACCTGGACATGGAGGCCACACAGTTGTGGCAGATCATCGTTCCTTTCTTATATGC 1270
 QY 1261 ATAAGATAACCAAGTGCATTTTATTTTTCGGCAGATTTCTCACCCTAAACCTAAGCGT 1320
 Db 1271 ATAAGATAACCAAGTGCATTTTATTTTTCGGCAGATTTCTCACCCTAAACCTAAGCGT 1330
 QY 1321 GGTGCTTCTGCAAAAGATTTTCTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAAT 1380
 Db 1331 GGTGCTTCTGCAAAAGATTTTCTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAAT 1390
 QY 1381 TGCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAACAACCTTCTGCTACC 1440
 Db 1391 TGCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAACAACCTTCTGCTACC 1450
 QY 1441 CACTAAATAAAACACAGAAATTTAGCAATTTCTATATATATATATATATATATATATAT 1500
 Db 1451 CACTAAATAAAACACAGAAATTTAGCAATTTCTATATATATATATATATATATATATATAT 1510
 QY 1501 AATCAATTTGGTCTTCTAAATGGGATCATGCCATTTAGATTTTCTTACTATCAGTTTA 1560
 Db 1511 AATCAATTTGGTCTTCTAAATGGGATCATGCCATTTAGATTTTCTTACTATCAGTTTA 1570
 QY 1561 TTTTATACATTAACCTTTTACTTTGTTATTTATTTATATATATATATATATATATATATAT 1620
 Db 1571 TTTTATACATTAACCTTTTACTTTGTTATTTATTTATATATATATATATATATATATATATAT 1630
 QY 1621 ATATATGCTCAGTCCCTTAATTAATGTAGCTAATAAAGTTATAGAGCAGATGATCTGTTA 1680
 Db 1631 ATATATGCTCAGTCCCTTAATTAATGTAGCTAATAAAGTTATAGAGCAGATGATCTGTTA 1690
 QY 1681 ATTTCTCTCTAATAAGCTTTTAAATTTCTCATATGAAGATAAGTAGGTATCCCT 1740
 Db 1691 ATTTCTCTCTAATAAGCTTTTAAATTTCTCATATGAAGATAAGTAGGTATCCCT 1749
 QY 1741 CCATGCCCTTCTGTAATAATATCTGGAATAAAACATTAAGCAATATATATATATATATAT 1800
 Db 1750 CCATGCCCTTCTGTAATAATATCTGGAATAAAACATTAAGCAATATATATATATATATAT 1809
 QY 1801 TGTGCAATTTCTAGAAATACATAACATATATATGCTGTATCTTATATATCAATTTCAAG 1860
 Db 1810 TGTGCAATTTCTAGAAATACATAACATATATATATGCTGTATCTTATATATCAATTTCAAG 1869
 QY 1861 TATATAATAATAAACCTGCTTCCAAACAAC 1891
 Db 1870 TATATAATAATAAACCTGCTTCCAAACAAC 1900

RESULT 6
 AAN70474 standard; cDNA; 2409 BP.
 ID AAN70474
 XX AAN70474;

XX 25-MAR-2003 (revised)
 DT 13-MAY-1991 (first entry)

XX Sequence encoding human minactivin.

XX Tumour location; tumour inhibition; chronic inflammation;
 XX rheumatoid arthritis; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 PH CDS 49..1296
 FT /*tag= a
 XX EP238275-A.
 XX 23-SEP-1987.
 XX 13-MAR-1987; 87EP-00302200.
 XX 13-MAR-1986; 86AU-00005017.
 PR 17-MAR-1986; 87AU-00071655.
 PR 22-MAY-1986; 86AU-00006033.
 PR 18-SEP-1986; 86AU-00008100.
 PR 21-NOV-1986; 86AU-00009104.
 XX (BIOT-) BIOTECHN AUSTR PTY.
 PA (AUSU) AUSTR NAT UNIV.
 PA Antalis TM, Barnes TM, Clark MA, Devine PL, Goss NH, Lehrbach PR;
 PI WPI; 1987-265615/38.
 DR P-PSDB; AAP70300.
 XX DNA sequence encoding minactivin - used for monitoring and treating
 PT tumours and chronic inflammation such as rheumatoid arthritis.
 XX Claim 3; pp34-37; 76pp; English.
 XX Minactivin is produced by induced human monocytes, certain macrophages
 CC and transformed cells of monocytic lineage (see WO 86/01212). Using the
 CC AA sequence obtd. for peptides of minactivin, oligonucleotide probes can
 CC be synthesised. (see AAN70467- AAN70473, and AAN71181-N71183). The
 CC specific oligonucleotide probe may be radiolabelled and then used to
 CC screen cDNA libraries to identify clones contg. all or part of the
 CC minactivin gene. (Updated on 25-MAR-2003 to correct PI field.)
 XX Sequence 2409 BP; 743 A; 500 C; 508 G; 658 T; 0 U; 0 Other;

Query Match 96.8%; Score 1846; DB 1; Length 2409;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1849; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 15 GTCCAGACAGCAACTCAGAGAAATACAGAGAAACACAGAGATTTGAACAATCGAGATCTT 74
 Db 1 GTCCAGACAGCAACTCAGAGAAATACAGAGAAACACAGAGATTTGAACAATCGAGATCTT 60
 QY 75 TGTGTGGCAACACACTCTTTGCCCTCAATTATTCAAGCATCTGGCAAAAGCAAGCCCC 134
 Db 61 TGTGTGGCAACACACTCTTTGCCCTCAATTATTCAAGCATCTGGCAAAAGCAAGCCCC 120
 QY 135 ACCCAGAACCTTCTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTACATG 194
 Db 121 ACCCAGAACCTTCTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTACATG 180
 QY 195 GGCTCCAGGGGAGCAGCAGAGCCAGATGGCCAGGTCCTTCAAGTTAATGAAGTGGGA 254
 Db 181 GGCTCCAGGGGAGCAGCAGAGCCAGATGGCCAGGTCCTTCAAGTTAATGAAGTGGGA 240
 QY 255 GCCAATGCAGTTACCCCATGACTCCAGAGAACCTTTACAGCTGTGGTTCATGCGACAG 314
 Db 241 GCCAATGCAGTTACCCCATGACTCCAGAGAACCTTTACAGCTGTGGTTCATGCGACAG 300
 QY 315 ATCCAGAGGGGTAGTTATCTCTGATGCGATTTTGCAGGCACCAAGCTGCAGATAAAATCCAT 374
 Db 301 ATCCAGAGGGGTAGTTATCTCTGATGCGATTTTGCAGGCACCAAGCTGCAGATAAAATCCAT 360
 QY 375 TCATCCTTCGGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTTACTGGAA 434
 Db 361 TCATCCTTCGGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTTACTGGAA 420
 QY 435 AGTGCTCAATAGCTGTTTGGTGAGAGTCTCGAGCTTCCGGGAAGAATATATTCGACTC 494

Db 421 AGTGCAATAGCTGTTGGTGAGAGTCTGCGAGCTTCGGGAGAGATATATTCGACTC 480
 Qy 495 TGTGAGAAATATATCTCTCAGAACCCGAGGAGTAGACTTCCTAGAAATGCGAGAGAA 554
 Db 481 TGTGAGAAATATATCTCTCAGAACCCGAGGAGTAGACTTCCTAGAAATGCGAGAGAA 540
 Qy 555 GCTAGAAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTG 614
 Db 541 GCTAGAAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTG 600
 Qy 615 TTACTGAGAGTCTGTAGATGGGATACAGAGATGCTGCTGGTGAATGCTGTACTTC 674
 Db 601 TTACTGAGAGTCTGTAGATGGGATACAGAGATGCTGCTGGTGAATGCTGTACTTC 660
 Qy 675 AAAGGAAAGTGGAAAACTCCATTTGAGAGAAACTAAATGGGCTTTATPCTTTCCGTGTA 734
 Db 661 AAAGGAAAGTGGAAAACTCCATTTGAGAGAAACTAAATGGGCTTTATPCTTTCCGTGTA 720
 Qy 735 AACTCGGCTCAGCGCACACCTGTACAGATGATGCTGGTGAAGAGCTAAACATTGGA 794
 Db 721 AACTCGGCTCAGCGCACACCTGTACAGATGATGCTGGTGAAGAGCTAAACATTGGA 780
 Qy 795 TACATAGAGACCTAAAGGCTCAGATTTAGAACTCCCATATGCTGGAGATGTTAGCATG 854
 Db 781 TACATAGAGACCTAAAGGCTCAGATTTAGAACTCCCATATGCTGGAGATGTTAGCATG 840
 Qy 855 TTCTTCTGCTCCAGATGAAATTTGCCGATGTGCTCCACTGGCTTGGAGCTGCTGGAAGT 914
 Db 841 TTCTTCTGCTCCAGATGAAATTTGCCGATGTGCTCCACTGGCTTGGAGCTGCTGGAAGT 900
 Qy 915 GAAATACCTATGACAACTCAACAGTGGACCGAGAAAGACAAATGGCTGGAAGATGA 974
 Db 901 GAAATACCTATGACAACTCAACAGTGGACCGAGAAAGACAAATGGCTGGAAGATGA 960
 Qy 975 GTTGAGGTATACATACCCAGTTCAAAATTAGAAGCAATTAAGACTCAGATCCCATTTCTG 1034
 Db 961 GTTGAGGTATACATACCCAGTTCAAAATTAGAAGCAATTAAGACTCAGATCCCATTTCTG 1020
 Qy 1035 AAAGCATGGGCTGAGGAGAGCCCTTCAAGAGGACGGGCAATTTCTCAGGATGTCG 1094
 Db 1021 AGAAGCATGGGCTGAGGAGAGCCCTTCAAGAGGACGGGCAATTTCTCAGGATGTCG 1080
 Qy 1095 GAGAGGAATGACCTGTTCTTCTTCTGAGTGTTCACCAAGGCAATGCTGGATGTAATGAG 1154
 Db 1081 GAGAGGAATGACCTGTTCTTCTTCTGAGTGTTCACCAAGGCAATGCTGGATGTAATGAG 1140
 Qy 1155 GAGGCACTGAGCAGCGCTGGACAGAGGTGTATGACAGGAGAGCTGACATGGA 1214
 Db 1141 GAGGCACTGAGCAGCGCTGGACAGAGGTGTATGACAGGAGAGCTGACATGGA 1200
 Qy 1215 GGCCACAGTTTGTGGCAGATCATCCGTTCTTTTCTTATTTATGATGATAGATAACCAAG 1274
 Db 1201 GGCCACAGTTTGTGGCAGATCATCCGTTCTTTTCTTATTTATGATGATAGATAACCAAC 1260
 Qy 1275 TGCAATTTATTTTGGCAGATTTGCTCACCTAAACTAGCGTCTCTGCAAA 1334
 Db 1261 TGCAATTTATTTTGGCAGATTTTCTCACCTAAACTAGCGTCTCTGCAAA 1320
 Qy 1335 AGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTCAAAATGCAAAATTTAG 1394
 Db 1321 AGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTCAAAATGCAAAATTTAG 1380
 Qy 1395 AGATGTTTCTACATATTTCTGCTCTTCTGACAACTTCTGCTACCCACTAAATAAAAC 1454
 Db 1381 AGATGTTTCTACATATTTCTGCTCTTCTGACAACTTCTGCTACCCACTAAATAAAAC 1440
 Qy 1455 ACAGAAATTAATAGACAAATGCTATTAATAACATGACACCCATTAATTAATTTGCTTT 1514
 Db 1441 ACAGAAATTAATAGACAAATGCTATTAATAACATGACACCCATTAATTAATTTGCTTT 1500
 Qy 1515 CTAAATGGGATCATGCCCATTTAGATTTCTTACTATCAGTTTATTTTATAACATTA 1574

Db 1501 CTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATTTTATAACATTA 1560
 Qy 1575 ACTTTTACTTTCTTATTTATTTATTTATATAAATGGTGAGTTTAAAAATTTATGCTCACTG 1634
 Db 1561 ACTTTTACTTTCTTATTTATTTATTTATATAAATGGTGAGTTTAAAAATTTATGCTCACTG 1620
 Qy 1635 CTTATTTATGATGCTAAATAAGTTATAGAGATGATCTGTTAAATTTCTTATCTCTAAT 1694
 Db 1621 CTTATTTATGATGCTAAATAAGTTATAGAGATGATCTGTTAAATTTCTTATCTCTAAT 1680
 Qy 1695 AAATGCTTTTAAATTTCTCATATGAAGAAATAGTAGGTATCCCTCCATCCCTTCTGT 1754
 Db 1681 AAATGCTTTTAAATTTCTCATATGAAGAAATAGTAGGTATCCCTCCATCCCTTCTGT 1740
 Qy 1755 AATAATATCTGGAAAAACATTAACATAGCAATATATGTTATGTCATTTCTTAGA 1814
 Db 1741 AATAATATCTGGAAAAACATTAACATAGCAATATATGTTATGTCATTTCTTAGA 1800
 Qy 1815 AATACATAACACATATATATGCTGTATCTTATTTCAATTTGCAAGTATATAAT 1869
 Db 1801 AATACATAACACATATATATGCTGTATCTTATTTCAATTTGCAAGTATATAAT 1854

RESULT 7
 AAN91177
 ID AAN91177 standard; DNA; 2424 BP.
 XX
 AC AAN91177;
 XX
 DT 25-MAR-2003 (revised)
 DT 10-MAR-2003 (revised)
 DT 12-JUN-1989 (first entry)
 XX
 DE Sequence encoding plasminogen activator inhibitor, type 2 (PAI-2).
 XX
 KW Plasminogen activator inhibitor; type 2; antibodies to PAI-2; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 553..1797
 FT polyA_signal /*tag= a
 FT 2382..2424 /*tag= b
 FT
 XX
 PN DE3722673-A.
 XX
 PD 19-JAN-1989.
 XX
 PF 09-JUL-1987; 87DE-03722673.
 XX
 PR 09-JUL-1987; 87DE-00713272.
 XX
 PA (BEHW) BEHRINGWERKE AG.
 XX
 PI Tor N, Lecander I, Astedt B;
 XX
 DR WPI; 1989-024876/04.
 DR P-PSDB; AAP94160.
 XX
 PT New DNA encoding complete plasminogen activator inhibitor type 2 -
 PT expressed protein and derived antibodies, useful therapeutically and
 PT diagnostically.
 XX
 PS Disclosure; Page 3-4; 4pp; German.
 XX
 CC Sequence encodes plasminogen activator inhibitor type 2 (PAI-2) which can
 CC be produced using recombinant DNA techniques. Pure PAI-2 can be produced
 CC in this way and has therapeutic uses. Antibodies to this protein can be
 CC used to diagnose PAI-2 genetic defects. Patent no. DE-3713272 described
 CC the 361 N-terminal AA's of this protein which are used here as probes to
 CC rescreen the original cDNA bank, allowing the new sequence, encoding the
 CC entire PAI-2 mol.to be isolated. This PAI differs from human placental

CC	PAI at posns. 120 (Asn replacing Asp), 404 (Asn replacing Lys), and 414 (Ser replacing Cys). In addition Proline at posn. 393 is encoded by CCT instead of CCG. (Updated on 10-MAR-2003 to add missing OS field.)
CC	(Updated on 25-MAR-2003 to correct FI field.)
XX	
SQ	Sequence 2424 BP; 723 A; 516 C; 533 G; 652 T; 0 U; 0 Other;
Query March	91.7%; Score 1750; DB 1; Length 2424;
Best Local Similarity	96.6%; Pred. No. 0;
Matches 1823; Conservative	0; Mismatches 5; Indels 60; Gaps 1;
QY	19 GACAGCAACTCAGAGATAACACAGACACACAGATTGAAACATGAGGATCTTTG 78
DB	569 GACAGCAACTCAGAGATAACACAGACACACAGATTGAAACATGAGGATCTTTG 628
QY	79 TGGCAACACACTCTTTGCCCTCAATTATTTCAAGCATCTGGCAAAAGAACCCACCC 138
DB	629 TGGCAACACACTCTTTGCCCTCAATTATTTCAAGCATCTGGCAAAAGAACCCACCC 688
QY	139 AGAACCTCTTCTCCCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTACATGGGCT 198
DB	689 AGAACCTCTTCTCCCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTACATGGGCT 748
QY	199 CCAGGGCAGCACCGAAGACAGATGGCAAGTGCTTCAGTTTAAATGAAGTGGAGCCA 258
DB	749 CCAGGGCAGCACCGAAGACAGATGGCAAGTGCTTCAGTTTAAATGAAGTGGAGCCA 808
QY	259 ATGCAGTTTACCCCATGACTCCAGAGAACTTTACCAGCTGTGGTTCATGCAGCAGATCC 318
DB	809 ATGCAGTTTACCCCATGACTCCAGAGAACTTTACCAGCTGTGGTTCATGCAGCAGATCC 868
QY	319 AGAAGGTAGTTATCTCGATGGATTTTGCAGGCAAGCTCCAGTAAATCCATTCAT 378
DB	869 AGAAGGTAGTTATCTCGATGGATTTTGCAGGCAAGCTCCAGTAAATCCATTCAT 928
QY	379 CTTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAAATTTACTGGAAAGTG 438
DB	929 CTTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAAATTTACTGGAAAGTG 988
QY	439 TCATAAGCTTTTGGTGGAGAGCTGGAGCTTCGGGAAGAAATATATTCGACTCTGTC 498
DB	989 TCATAAGCTTTTGGTGGAGAGCTGGAGC-----1020
QY	499 AGAAATATTAATCTCTCAGAACCCAGGCAAGTAGCTTCTAGATGTGCAGAGAAGCTTA 558
DB	1021 -----GACTTCTAGAAATGTGCAGAGAAGCTTA 1048
QY	559 GAAABAGATTAAATCTCGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTGTATC 618
DB	1049 GAAABAGATTAAATCTCGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTGTATC 1108
QY	619 CTGAAGGTTCTGTAGATGGGATACAGGATGGTCTGGTGAATGTGTCTACTTCADAG 678
DB	1109 CTGAAGGTTCTGTAGATGGGATACAGGATGGTCTGGTGAATGTGTCTACTTCADAG 1168
QY	679 GAAAGTGGAAACTCCATTGTAGAGAAACTAAATGGGCTTTATCTTTCCGTGTAACT 738
DB	1169 GAAAGTGGAAACTCCATTGTAGAGAAACTAAATGGGCTTTATCTTTCCGTGTAACT 1228
QY	739 CGGCTCAGGCGCACCTGTACAGATGATGTACTTGGTGAAGCTTAAACATTGGATACA 798
DB	1229 CGGCTCAGGCGCACCTGTACAGATGATGTACTTGGTGAAGCTTAAACATTGGATACA 1288
QY	799 TAGAAGACCTTAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGATGTTAGCATGTTCT 858
DB	1289 TAGAAGACCTTAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGATGTTAGCATGTTCT 1348
QY	859 TGTTCCTCCAGATGAATTCGGATGTGTCCTACTGGCTGGAGCTGCTGGAAGTGA 918
DB	1349 TGTTCCTCCAGATGAATTCGGATGTGTCCTACTGGCTGGAGCTGCTGGAAGTGA 1408
QY	919 TAACTTATGACAACTCAACAGTGGACCGCAAGACAAATGGCTGGAAGTGAAGTTG 978

RESULT 8
AAN80060
ID AAN80060 standard; cDNA; 1675 BP.
XX
AC AAN80060;
XX

1409	TAACTTATGACAACTCAACAGTGGACCAAGACAAATGGCTGAAGATGAAGTTG 1468
979	AGTATACATACCCAGTTCAAATTAAGAGAGATTATGAATCAGATCCATTTCTGAAA 1038
1469	AGTATACATACCCAGTTCAAATTAAGAGAGATTATGAATCAGATCCATTTCTGAGAA 1528
1039	GAATGGCATGAGAGAGCCCTTCAACAGGAGCGGCCAATTTCTCAGGATGTCGGAGA 1098
1529	GCATGGCATGAGAGAGCCCTTCAACAGGAGCGGCCAATTTCTCAGGATGTCGGAGA 1588
1099	GAATGACCTGTTTCTTTCTGAAAGTTTCCACCAAGCCATGGTGGATGTAATGAGGAG 1158
1589	GAATGACCTGTTTCTTTCTGAAAGTTTCCACCAAGCCATGGTGGATGTAATGAGGAG 1648
1159	GCATCAAGCAGCGCTGGCAGAGAGGTGTATGACAGGAGAGAACTGACATGAGGCGC 1218
1649	GCATCAAGCAGCGCTGGCAGAGAGGTGTATGACAGGAGAGAACTGACATGAGGCGC 1708
1219	CACAGTTTGGCAGATCATCGGTTCTTTCTTTATTTATGATGAATGAATGAATGCA 1278
1709	CACAGTTTGGCAGATCATCGGTTCTTTCTTTATTTATGATGAATGAATGAATGCA 1768
1279	TTTTTTTTCGGCAGATTTTCTCAACCTAAACCTAAGCGTGTCTTCTGCAAAAGAT 1338
1769	TTTTTTTTCGGCAGATTTTCTCAACCTAAGCGTGTCTTCTGCAAAAGAT 1828
1339	TTTTGTAGATGAGCTGTGCTCAGAAATTCCTATTTCAATTTGCCAAAATTTAGAGAT 1398
1829	TTTTGTAGATGAGCTGTGCTCAGAAATTCCTATTTCAATTTGCCAAAATTTAGAGAT 1888
1399	GTTTTCTACATATTTCTGCTCTTCTGAACAACTTCTGCTACCCACTAAATAAAAAACAG 1458
1889	GTTTTCTACATATTTCTGCTCTTCTGAACAACTTCTGCTACCCACTAAATAAAAAACAG 1948
1459	AAATAATAGACATTTCTTATTAATGATGAGAGCTTATTAATCAATCAATTTGCTCTTAA 1518
1949	AAATAATAGACATTTCTTATTAATGATGAGAGCTTATTAATCAATCAATTTGCTCTTAA 2008
1519	AATGGGATCATGCCCATTTAGATTTCTTACTATCATGTTTATTTTATAACATTAACCTT 1578
2009	AATGGGATCATGCCCATTTAGATTTCTTACTATCATGTTTATTTTATAACATTAACCTT 2068
1579	TTACTTTGTTATTTATTTATTAATGATGAGTTTATAATTTTAAATTTATGCTCACTGCTTA 1638
2069	TTACTTTGTTATTTATTTATTAATGATGAGTTTATAATTTTAAATTTATGCTCACTGCTTA 2128
1639	TTTAAATGATGATAATAAGTTATAGAGCAGATGATCTGTTAATTTCTCTATCTTAATAAT 1698
2129	TTTAAATGATGATAATAAGTTATAGAGCAGATGATCTGTTAATTTCTCTATCTTAATAAT 2188
1699	GCCTTTAATTTGTTCTCAATAAGAGAAATAAGTAGGTATCCCTCCATGCCCTTCTGTAATA 1758
2189	GCCTTTAATTTGTTCTCAATAAGAGAAATAAGTAGGTATCCCTCCATGCCCTTCTGTAATA 2248
1759	AATATCTGGAAAAACATTAACATAGCAATATATGTTATGTCATTTCTAGAAATA 1818
2249	AATATCTGGAAAAACATTAACATAGCAATATATGTTATGTCATTTCTAGAAATA 2308
1819	CATAACACATATATGTCGTCTTATTTCAATTTCAATTTCAAGATGATATAATAATAACCT 1878
2309	CATAACACATATATGTCGTCTTATTTCAATTTCAATTTCAAGATGATATAATAATAACCT 2368
1879	GCCTTCCAAACCAACAAAAATAAAAAAA 1906
2369	GCCTTCCAAACCAACAAAAATAAAAAAA 2396

05-NOV-1990 (first entry)
Sequence encoding plasminogen activator inhibitor type 2 (PAI2) in human
placental cDNA clone lambda-plc 1.
Protease inhibitor; thrombolytic; ss.
Homo sapiens.
Key Location/Qualifiers
CDS 1..1086
/*tag= a
DE3713272-A.
03-NOV-1988.
18-APR-1987; 87DE-03713272.
09-JUL-1987; 87DE-03722673.
(BEHW) BEHRINGWERKE AG.
Ny T, Lecander I, Astedt B;
WPI; 1988-315694/45.
P-PSDB; AAP80058.
New DNA sequence encoding plasminogen activator inhibitor type 2 - and
derived proteins and antibodies, useful therapeutically and in diagnosis.
Claim 1; Table 1 Page 5; 6pp; German.
Commercially available human placental cDNA bank in lambda gt11 phage was
used to transform E.coli Y1090 and culture supernatants from the
transfomants screened by reaction with mouse Ab. against PAI2. Clone
lambda-plc 1 was one of 6 clones isolated. PAI2 is a strong and specific
inhibitor for the active forms of PA so is used to treat or prevent
disorders associated with abnormal levels of PA. PA is implicated in
haemostasis, tissue repair, metastasis formation, ovulation, fertility,
macrophage migration and inflammatory processes
Sequence 1675 BP; 530 A; 322 C; 331 G; 492 T; 0 U; 0 Other;
Query Match 87.4%; Score 1667; DB 1; Length 1675;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1670; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 225 GCCAAGGTCCTCAGTTAATGAAGTGGAGCCAAATCAGTTACCCCATGACTCCAGAG 284
DB 1 GCCAAGGTCCTCAGTTAATGAAGTGGAGCCAAATCAGTTACCCCATGACTCCAGAG 60
QY 285 AACTTTACCAGCTGTGGGTTTCATGCAGCAGATCCAGAAGGGTAGTTATCCTGATGCGATT 344
DB 61 AACTTTACCAGCTGTGGGTTTCATGCAGCAGATCCAGAAGGGTAGTTATCCTGATGCGATT 120
QY 345 TTGCAGGCAACAGCTGCAGATAAATCCATTCCTTCGCTCTCAGCTCTGCAATC 404
DB 121 TTGCAGGCAACAGCTGCAGATAAATCCATTCCTTCGCTCTCAGCTCTGCAATC 180
QY 405 AATGCATCCACAGGGAATTTACTGGAAGTGTCAATGAAGCTGTTTGGTGAGAGTCT 464
DB 181 AATGCATCCACAGGGAATTTACTGGAAGTGTCAATGAAGCTGTTTGGTGAGAGTCT 240
QY 465 GCGAGCTTCGGAAGATATATTCGACTGTGCAGAAATATTACTCTCAGAACCCOCAG 524
DB 241 GCGAGCTTCGGAAGATATATTCGACTGTGCAGAAATATTACTCTCAGAACCCOCAG 300
QY 525 GCAGTAGACTTCTAGAAATGTGCAGAAAGAGCTAGAAAAAGATTAATTCCTGGGTCAAG 584
DB 301 GCAGTAGACTTCTAGAAATGTGCAGAAAGAGCTAGAAAAAGATTAATTCCTGGGTCAAG 360
QY 585 ACTCAACCAAGGCAAAATCCCAAACTTGTTACCTGAAGTTCTGTAGATGGGATACC 644

Db 361 ACTCAACCAAGGCAAAATCCCAAACTTGTTTACTGAAGTTCTGTAGATGGGATACC 420
QY 645 AGGATGTCCTGGTGAATGCTGTCTACTTCAAAGGAAAGTGGAAACTCCATTTGAGAAG 704
Db 421 AGGATGTCCTGGTGAATGCTGTCTACTTCAAAGGAAAGTGGAAACTCCATTTGAGAAG 480
QY 705 AAACCTAAATGGGCTTTATCCTTTCCGTGTAAACTCGGCTCAGCGCACACCTGTACAGATG 764
Db 481 AAACCTAAATGGGCTTTATCCTTTCCGTGTAAACTCGGCTCAGCGCACACCTGTACAGATG 540
QY 765 ATGTACTTGGGTGAAAGCTAAACATTTGATATAGAAAGACCTTAAGGCTCAGATTCTA 824
Db 541 ATGTACTTGGGTGAAAGCTAAACATTTGATATAGAAAGACCTTAAGGCTCAGATTCTA 600
QY 825 GAACTCCCATATGCTGGAGATGTTAGCATGTTCTTGTTCAGATGAAATTCGCGAT 884
Db 601 GAACTCCCATATGCTGGAGATGTTAGCATGTTCTTGTTCAGATGAAATTCGCGAT 660
QY 885 GTGTCCATGCTTGGAGCTGTGAAAGTGAATTAACCTATGACAAACTCAACAAGTGG 944
Db 661 GTGTCCATGCTTGGAGCTGTGAAAGTGAATTAACCTATGACAAACTCAACAAGTGG 720
QY 945 ACCGACAAAGACAAAATGGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTA 1004
Db 721 ACCGACAAAGACAAAATGGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTA 780
QY 1005 GAAGAGCATTTATGAATCTAGATCCATTCTGAAAGCATGGGCATGAGGAGCCCTTCAAC 1064
Db 781 GAAGAGCATTTATGAATCTAGATCCATTCTGAAAGCATGGGCATGAGGAGCCCTTCAAC 840
QY 1065 AAGGACCGGCGCAATTTCTCAGGCGATGTCGAGAGGAATGACCTGTTTCTTCTGAAGTG 1124
Db 841 AAGGACCGGCGCAATTTCTCAGGCGATGTCGAGAGGAATGACCTGTTTCTTCTGAAGTG 900
QY 1125 TTCCACCAAGCCATGTTGATGATGAGGAGGCACTGAAGCAGCGCTGGCACAGGA 1184
Db 901 TTCCACCAAGCCATGTTGATGATGAGGAGGCACTGAAGCAGCGCTGGCACAGGA 960
QY 1185 GGTGTTATGACAGGAGAACTGGACATGGAGGCCACAGTTTGTGGCAGATCATCCGTTT 1244
Db 961 GGTGTTATGACAGGAGAACTGGACATGGAGGCCACAGTTTGTGGCAGATCATCCCTTT 1020
QY 1245 CTTTTTCTTATTTATGATTAAGATAACCAAGTGCATTTTATTTTCGGCAGATTTTGTCTCA 1304
Db 1021 CTTTTTCTTATTTATGATTAAGATAACCAAGTGCATTTTATTTTCGGCAGATTTTCTCTCA 1080
QY 1305 CCCTAAAACTAAGCGTGTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAG 1364
Db 1081 CCCTAAAACTAAGCGTGTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAG 1140
QY 1365 AATTGCTATTTCAAAATGGCAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTG 1424
Db 1141 AATTGCTATTTCAAAATGGCAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTG 1200
QY 1425 AACCAACTTCTGCTTACCCTAAATAAAAAACACAGAAATAAATTAGACAATTTCTATTATA 1484
Db 1201 AACCAACTTCTGCTTACCCTAAATAAAAAACACAGAAATAAATTAGACAATTTCTATTATA 1260
QY 1485 ACATGACAACTTATTAATCAATTTGGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1544
Db 1261 ACATGACAACTTATTAATCAATTTGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1320
QY 1545 CCTTACTATCAGTTTATTTTATTAACATTAACCTTTTACTTTTCTTCTTCTTCTTCTTCTT 1604
Db 1321 CCTTACTATCAGTTTATTTTATTAACATTAACCTTTTACTTTTCTTCTTCTTCTTCTTCTT 1380
QY 1605 AATGGTGAAGTTTAAATATTTCTGCTCCTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1664
Db 1381 AATGGTGAAGTTTAAATATTTCTGCTCCTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
QY 1665 AGCAGATGATCTGTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1724

Db 1441 AGCAGATGCTGTTAAATTCCTATCTATAAATGCTTTAAATGTTCTCATATGAAGA 1500
Qy 1725 ATAAGTAGTATCCCTCCATGCCCTTCTGTAATAAATATCTGAAATAAATTAACAAT 1784
Db 1501 ATAAGTAGTATCCCTCCATGCCCTTCTGTAATAAATATCTGAAATAAATTAACAAT 1560
Qy 1785 AGGCAATATATGTTATGTCATTTCTAGAAATACATAACACATATATATGTCGTATCT 1844
Db 1561 AGGCAATATATGTTATGTCATTTCTAGAAATACATAACACATATATATGTCGTATCT 1620
Qy 1845 TATATTCATTCGACATATATAAATAAATACCTGCTTCCAAACACAAAAATA 1899
Db 1621 TATATTCATTCGACATATATAAATAAATACCTGCTTCCAAACACAAAAATA 1675

RESULT 9
AAQ10649
ID AAQ10849 standard; DNA; 1412 BP.
XX AC AAQ10849;
XX AC
XX DT 25-MAR-2003 (revised)
XX DT 08-MAY-1991 (first entry)
XX DE Encodes Plasminogen Activator Inhibitor-2.
XX KW plasminogen activator inhibitor-2; PAI-2; inflammation treatment; cancer;
XX KW ss.
XX OS Homo sapiens.
XX FH
XX FT Key Location/Qualifiers
XX FT CDS 42..1289
XX FT /*tag= a
XX FT /product= "PAI-2"
XX FN
XX PD WO9102057-A.
XX PD 21-FEB-1991.
XX PF 09-AUG-1989; 89GB-00018191.
XX PF 09-AUG-1989; 89GB-00018191.
XX PF 05-DEC-1989; 89GB-00027480.
XX PF (DELZ) DELTA BIOTECHNOLOGY LTD.
XX PI
XX PI Ballance DJ, Steven J;
XX DR WPI; 1991-073533/10.
XX DR P-PSDB; AAR10921.
XX PT Prepn. of plasminogen activator inhibitor-2 - by fermentation of a fungal
XX PT cell transformed with a genetic construct which expresses the plasminogen
XX PT activator inhibitor-2.
XX PS
XX PS Example; Fig 1; 52pp; English.
XX CC A lambda gt11 cDNA library constructed from phorbol-12-myristate-13-
XX CC acetate stimulated cells of the human monocyte-like histiocytic lymphoma
XX CC cell line U937 was screened. (See AAQ10850 and AAQ10851 for the probe
XX CC sequences). From the putative positive clones, one was selected (lambda
XX CC gt11-186), which appeared to contain the entire PAI-2 coding region. The
XX CC insert was cloned in M13mp19 to give pDBP1 and was sequenced. (Updated on
XX CC 25-MAR-2003 to correct PA field.)
XX SQ
XX SQ Sequence 1412 BP; 418 A; 311 C; 323 G; 360 T; 0 U; 0 Other;
Query Match 73.7%; Score 1405.6; DB 2; Length 1412;
Best Local Similarity 99.7%; Pred. No. 3.1e-281;
Matches 1408; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 22 AGCAACTCAGAGATACACCAACACCAACGATTCGAGGATCTTGTGTGG 81

Db 1 AGCACTCAGAGATACACCAAGAGACCAACAGATTGAAACAATGGAGATCTTTGTGTGG 60
Qy 82 CAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCAACCCAGA 141
Db 61 CAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCAACCCAGA 120
Qy 142 ACCTCTTCTCTCCCATCGAGCATCTCGTCCACCATGGCCATGGCTCTACATGGGCTCA 201
Db 121 ACCTCTTCTCTCCCATCGAGCATCTCGTCCACCATGGCCATGGCTCTACATGGGCTCA 180
Qy 202 GGGGAGAGCCGAGACGAGATGGCCAAAGTGCTTCAAGTTTAATGAAGTGGAGCCCAATG 261
Db 181 GGGGAGAGCCGAGACGAGATGGCCAAAGTGCTTCAAGTTTAATGAAGTGGAGCCCAATG 240
Qy 262 CAGTTACCCCATGACTCCAGAGAACTTTACAGCTGTGGGTTCATGGAGAGATCCAGA 321
Db 241 CAGTTACCCCATGACTCCAGAGAACTTTACAGCTGTGGGTTCATGGAGAGATCCAGA 300
Qy 322 AGGGTAGTTATCTGATGCAATTTTGCAGGCACAAGCTGCAGATAAAATCCATTCATCCT 381
Db 301 AGGGTAGTTATCTGATGCAATTTTGCAGGCACAAGCTGCAGATAAAATCCATTCATCCT 360
Qy 382 TCCGCTCTCTAGCTCTCAATCAATGCAATCCAGGGAATTTATTACTGGAAGTGTCA 441
Db 361 TCCGCTCTCTAGCTCTCAATCAATGCAATCCAGGGAATTTATTACTGGAAGTGTCA 420
Qy 442 ATAAGCTGTTTGTGAGAAAGTCTGCGAGCTTCCGGGAAGAAATATATTGCACTCTGTCA 501
Db 421 ATAAGCTGTTTGTGAGAAAGTCTGCGAGCTTCCGGGAAGAAATATATTGCACTCTGTCA 480
Qy 502 AATATTACTCTCAGAAACCCAGGAGTAGACTCTTAGATGTGCGAAGAGAGTAGAA 561
Db 481 AATATTACTCTCAGAAACCCAGGAGTAGACTCTTAGATGTGCGAAGAGAGTAGAA 540
Qy 562 AAAAGATTAAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCAACTTTTACCTG 621
Db 541 AAAAGATTAAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCAACTTTTACCTG 600
Qy 622 AAGGTTCTGTAGATGGGATACAGGATGCTCCTGGTGAATGCTGTCTACTTCAAGGAA 681
Db 601 AAGGTTCTGTAGATGGGATACAGGATGCTCCTGGTGAATGCTGTCTACTTCAAGGAA 660
Qy 682 AGTGAAAACTCCATTTGAGAGAAACTAAATGGGCTTTATCTCTTCCGTGTAATCTGG 741
Db 661 AGTGAAAACTCCATTTGAGAGAAACTAAATGGGCTTTATCTCTTCCGTGTAATCTGG 720
Qy 742 CTCAGCGCACCTGTACAGATGATGTTCTGCTGAAAGCTAAACATTTGGATACATAG 801
Db 721 CTCAGCGCACCTGTACAGATGATGTTCTGCTGAAAGCTAAACATTTGGATACATAG 780
Qy 802 AAGACCTAAAGGCTCAGATTTAGAACTCCCATATGCTGGAGATGTTAGCATGTTCTTGT 861
Db 781 AAGACCTAAAGGCTCAGATTTAGAACTCCCATATGCTGGAGATGTTAGCATGTTCTTGT 840
Qy 862 TGCTTCCAGATGAAATTTGCCGATGTGTCCTACTGGCTTGGAGCTGTGGAAGTGAATAA 921
Db 841 TGCTTCCAGATGAAATTTGCCGATGTGTCCTACTGGCTTGGAGCTGTGGAAGTGAATAA 900
Qy 922 CCTATGACAACTCAACAGTGGACCAAGCAAAAGCAAAATGCTCAAGATGAAGTTGAGG 981
Db 901 CCTATGACAACTCAACAGTGGACCAAGCAAAAGCAAAATGCTCAAGATGAAGTTGAGG 960
Qy 982 TATACATACCCCAAGTTCAAATTTAGAGAGCAATTAAGAACTCAGATCCATTTCTGAAAGCA 1041
Db 961 TATACATACCCCAAGTTCAAATTTAGAGAGCAATTAAGAACTCAGATCCATTTCTGAGAGCA 1020
Qy 1042 TGGGCATGGAGAGCGCTTCAACAGGAGCGGGCCAAATTTCTCAGGGATGTCGGAGAGCA 1101
Db 1021 TGGGCATGGAGAGCGCTTCAACAGGAGCGGGCCAAATTTCTCAGGGATGTCGGAGAGCA 1080
Qy 1102 ATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATCGTGGATGTGAATGAGGAGGCA 1161

Db 702 AGATGTTAGCATGTTCTTGTGCTTCCAGATGAATGCCGATGTGTCACATGCTGCTGGA 761
Qy 902 GCTGCTGGAAGTGAATACCTATACAAACTCACAAGTGGACCAAGCAAGCAAAAT 961
Db 762 GCTGCTGGAAGTGAATACCTATACAAACTCACAAGTGGACCAAGCAAGCAAAAT 821
Qy 962 GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAGAGCATTTGAAC 1021
Db 822 GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAGAGCATTTGAAC 881
Qy 1022 CAGATCCATTCTGAAAAGCATGGCATGGAGGCGCTTCAACAAGGAGCGGGCAATTT 1081
Db 882 CAGATCCATTCTGAAAAGCATGGCATGGAGGCGCTTCAACAAGGAGCGGGCAATTT 941
Qy 1082 CTCAGGATGTCGAGAGGATGACCTGTTTCTTCTGAAGTGTCCACCAAGCATGGT 1141
Db 942 CTCAGGATGTCGAGAGGATGACCTGTTTCTTCTGAAGTGTCCACCAAGCATGGT 1001
Qy 1142 GGATGTGAATGAGGAGGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGAG 1201
Db 1002 GGATGTGAATGAGGAGGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGAG 1061
Qy 1202 AACTGGACATGGAGGCCACAGTTTGGGAGATCATCCGTTCTTCTTATTATGCA 1261
Db 1062 AACTGGACATGGAGGCCACAGTTTGGGAGATCATCCGTTCTTCTTATTATGCA 1121
Qy 1262 TAAGTAACCAAGTGCATTTATTTTCGGCAGATTTTGTCTACCCATAAACTAAGCGTG 1321
Db 1122 TAAGTAACCAAGTGCATTTATTTTCGGCAGATTTTGTCTACCCATAAACTAAGCGTG 1181
Qy 1322 CTGCTTCTGAAAAGATTTTGTAGATGAGTGTGCTCAGAAATGCTATTTCAAAT 1381
Db 1182 CTGCTTCTGAAAAGATTTTGTAGATGAGTGTGCTCAGAAATGCTATTTCAAAT 1241
Qy 1382 GCAAAAATTTAGAGATTTTCTACATATTTCTGCTTCTGACAACTCTGCTACCC 1441
Db 1242 GCAAAAATTTAGAGATTTTCTACATATTTCTGCTTCTGACAACTCTGCTACCC 1301
Qy 1442 ACTAAATAAAAAACACAGAAATAATTAGACAAATTTCTATTATAACATGACAAACCTATTA 1501
Db 1302 ACTAAATAAAAAACACAGAAATAATTAGACAAATTTCTATTATAACATGACAAACCTATTA 1361
Qy 1502 ATCATTGCTCTCTAAAAGGGATCATGCCATTTAGATTTTCTTACTATCAGTTTAT 1561
Db 1362 ATCATTGCTCTCTAAAAGGGATCATGCCATTTAGATTTTCTTACTATCAGTTTAT 1421
Qy 1562 TTTTATAACATTAACCTTTTACTTTGTTATTTATTTATTTATTAATGTTGAGTTT 1618
Db 1422 TTTTATAACATTAACCTTTTACTTTGTTATTTATTTATTTATTAATGTTGAGTTT 1478

RESULT 12
ID AAQ11128
XX AAQ11128 standard; DNA; 1328 BP.

XX AC AAQ11128;
XX AC
DT 09-JAN-2003 (revised)
DT 06-JUN-1991 (first entry)

XX DE PreA-plasminogen activator inhibitor-2 with signal sequence.

XX KW Plasminogen activator inhibitor; PAI-2; signal sequence; diagnosis;
XX KW treatment; tumour; inflammation; secretion; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers
XX CDS 15..1316

FT FT /*tag= a
FT FT /product= "preB-PAI-2"
FT sig_peptide 15..71

FT FT /*tag= b
XX XX /note= "synthetic-claim3"
PN WO9103556-A.
XX PD 21-MAR-1991.
XX PF 05-SEP-1989; 89AU-00006179.
XX PR 05-SEP-1989; 89AU-00006179.
XX (BIOT-) BIOTECHN AUST PTY L.
XX Whitfield P, Richardson MA, Bunn CL;
XX WPI: 1991-102074/14.
XX P-PSDB; AAR11352.
XX Prepn. of glycosylated plasminogen activator inhibitor-2 (PAI-2) - for
PT diagnosis and treatment of tumours, and in treatment of
PT immuno-suppressant conditions and inflammation.
XX Disclosure; Fig 7; 70pp; English.
XX Addition of a signal sequence to PAI-2 not only facilitates secretion of
CC the glycosylated 60 kD form of PAI-2 but also directs correct processing
CC of the signal. The signal-contg. variant was constructed such that the
CC artificial signal was fused to the NH2-terminal methionine of PAI-2.
CC Correct cellular processing of this signal peptide should result in a
CC mature PAI-2 molecule of 415 amino acids with Met at its NH2-terminus.
CC See AAQ11129 for PAI-2 with the alpha-1-antitrypsin signal sequence.
XX (Updated on 09-JAN-2003 to add missing OS field.)
XX Sequence 1328 BP; 377 A; 300 C; 317 G; 334 T; 0 U; 0 Other;

Query Match 65.5%; Score 1249.4; DB 2; Length 1328;
Best Local Similarity 99.5%; Pred. No. 6.6e-249;
Matches 1253; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 61 CAATGAGATCTTTGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAGCATCTGG 120
Db 70 CTATGGAAGATCTTTGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAGCATCTGG 129
Qy 121 CAAAGCAAGCCGCCACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACATGG 180
Db 130 CAAAGCAAGCCGCCACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACATGG 189
Qy 181 CCATGCTCTACATGGCTCCAGGGGAGACCCGAGACCGAGATGCGCAAGGTGCTTCAGT 240
Db 190 CCATGCTCTACATGGCTCCAGGGGAGACCCGAGACCGAGATGCGCAAGGTGCTTCAGT 249
Qy 241 TTAATGAAGTGGAGCCCAATGCGAGTTACCCCATGATGCTCCAGAGAACTTTACCAAGTGTG 300
Db 250 TTAATGAAGTGGAGCCCAATGCGAGTTACCCCATGATGCTCCAGAGAACTTTACCAAGTGTG 309
Qy 301 GGTTCATGACGACAGATCCAGAAGGGTAGTTATCTGATGCCATTTTGAGGCACCAAGCTG 360
Db 310 GGTTCATGACGACAGATCCAGAAGGGTAGTTATCTGATGCCATTTTGAGGCACCAAGCTG 369
Qy 361 CAGATAAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATCCATCCAGGGA 420
Db 370 CAGATAAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATCCATCCAGGGA 429
Qy 421 ATTATTACTGGAAAGTGTCAATAAGCTGTTGTGTGAGAGTCTGCGAGTCTCCGGGAAG 480
Db 430 ATTATTACTGGAAAGTGTCAATAAGCTGTTGTGTGAGAGTCTGCGAGTCTCCGGGAAG 489
Qy 481 AATATATTCGACTCTGTGAGAAATATTACTCTCAGAACCCCGAGCAGTAGACTTCTCTAG 540
Db 490 AATATATTCGACTCTGTGAGAAATATTACTCTCAGAACCCCGAGCAGTAGACTTCTCTAG 549
Qy 541 AATGTGACAGAAAGCTAGAAAAAGATTAATTCCTGGTCAAGACTCAAAACCAAGGCA 600

550 AATGTCAGAGAGCTAGAAAAAGATTAAATCTCGGTCAAGACTCAAAACCAAGGCA 609
601 AATCCCAACCTGTTACCTGAGGTTCTGTAGATGGGTACACAGATGCTCTGGTGA 660
610 AATCCCAACCTGTTACCTGAGGTTCTGTAGATGGGTACACAGATGCTCTGGTGA 669
661 ATGCTGTCTACTTCAAGGAAAGTGGAAACTCCATTTGAGAGAACTAAATGGGCTTT 720
670 ATGCTGTCTACTTCAAGGAAAGTGGAAACTCCATTTGAGAGAACTAAATGGGCTTT 729
721 ATCCTTTCCGTGTAATCTCGGCTCAGCGCACCTGTACAGATGATGATGCTGTGAAA 780
730 ATCCTTTCCGTGTAATCTCGGCTCAGCGCACCTGTACAGATGATGATGCTGTGAAA 789
781 AGCTAAACATTCATATAGAGACCTAAGGCTCAGATTCAGATCCCATATGCTG 840
790 AGCTAAACATTCATATAGAGACCTAAGGCTCAGATTCAGATCCCATATGCTG 849
841 GAGATGTTAGCATGTTCTTGTGCTTCAGATGAAATGCGGATGTCCTGCTGG 900
850 GAGATGTTAGCATGTTCTTGTGCTTCAGATGAAATGCGGATGTCCTGCTGG 909
901 AGCTGTGGAAGTGAATTAACCTATGACAACTCAACAAGTGGACCAAGCAAGCAAAA 960
910 AGCTGTGGAAGTGAATTAACCTATGACAACTCAACAAGTGGACCAAGCAAGCAAAA 969
961 TGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATAGAGAGCATTTAGAC 1020
970 TGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATAGAGAGCATTTAGAC 1029
1021 TCAGATCCATCTTGAAGAGCATGGGCTATGGAGGAGCGCTTCAACAGAGGACGGGCAATT 1080
1030 TCAGATCCATCTTGAAGAGCATGGGCTATGGAGGAGCGCTTCAACAGAGGACGGGCAATT 1089
1081 TCTCAGGATGTCGGAGGAGTACCTGTTTCTTCTGAGATGTTTCCACCAAGCATGG 1140
1090 TCTCAGGATGTCGGAGGAGTACCTGTTTCTTCTGAGATGTTTCCACCAAGCATGG 1149
1141 TGGATGTAATGAGAGGAGCTAGAGCAGCGCTGGCAGAGGAGTGTATGACAGGGA 1200
1150 TGGATGTAATGAGAGGAGCTAGAGCAGCGCTGGCAGAGGAGTGTATGACAGGGA 1209
1201 GAATGACATGAGAGGAGCTAGAGTGTGGAGATCATCCGTTCTTTTCTTATATGC 1260
1210 GAATGACATGAGAGGAGCTAGAGTGTGGAGATCATCCCTTTCTTTTCTTATATGC 1269
1261 ATAAGATAACCAAGTGCATTTATTTTTCGGAGATTTTGTCTCACCCTAAACATGCG 1319
1270 ATAAGATAACCAAGTGCATTTATTTTTCGGAGATTTTCTCACCCTAAACATGCG 1328

RESULT 13
AAQ11129 standard; DNA; 1340 BP.
XX AC AAQ11129;
XX
DT 09-JAN-2003 (revised)
DT 06-JUN-1991 (first entry)
XX
DE PreB-PAI-2 with alpha-1-antitrypsin signal sequence.
XX
KW Plasminogen activator inhibitor; PAI-2; signal sequence; diagnosis;
KW treatment; tumour; inflammation; secretion; alpha-1-antitrypsin; ss.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FT CDS 15..1328
FT /*tag= a
FT /product= "preB-PAI-2"
FT sig_peptide 15..86
FT /*tag= b

/label= alpha-1-antitrypsin
WO9103556-A.
21-MAR-1991.
05-SEP-1989; 89AU-00006179.
05-SEP-1989; 89AU-00006179.
(BIOT-) BIOTECHN AUST PTY L.
Whitfield P, Richardson MA, Bunn CL;
WPI; 1991-102074/14.
P-PSDB; AAR11353.
Prepn. of glycosylated plasminogen activator inhibitor-2 (PAI-2) - for
diagnosis and treatment of tumours, and in treatment of
immuno-suppressant conditions and inflammation.
Disclosure; Fig 8; 70pp; English.
Addition of a signal sequence to PAI-2 not only facilitates secretion of
the glycosylated 60 kD form of PAI-2 but also directs correct processing
of the signal. The signal-contg. variant was constructed such that the
human alpha-1-antitrypsin signal was fused to PAI-2 without its NH2-
terminal methionine. Correct cellular processing of this signal peptide
should result in a mature PAI-2 molecule of 414 amino acids with Glu at
its NH2-terminus. See ARQ11128 for PAI-2 with a synthetic signal
sequence. (Updated on 09-JAN-2003 to add missing OS field.)
XX
SQ Sequence 1340 BP; 374 A; 307 C; 318 G; 341 T; 0 U; 0 Other;
Query Match 65.1%; Score 1242.8; DB 2; Length 1340;
Best Local Similarity 99.4%; Pred. No. 1.5e-247;
Matches 1247; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 66 GAGATCTTTTGTGTGGCAACACACTCTTTCGCCCTCAATTTATTTCAGCATCTGGCAAAA 125
DB 87 GAGATCTTTTGTGTGGCAACACACTCTTTCGCCCTCAATTTATTTCAGCATCTGGCAAAA 146
QY 126 GCAAGCCCCACCCAGAACCTTCTCTCCCATGGAGCATCTGTCACCATGGCCATG 185
DB 147 GCAAGCCCCACCCAGAACCTTCTCTCCCATGGAGCATCTGTCACCATGGCCATG 206
QY 186 GTCTACATGGCTCCAGGGGAGCAGCAGAACACCATGGCCAGGTGCTTCAGTTTAAAT 245
DB 207 GTCTACATGGCTCCAGGGGAGCAGCAGAACACCATGGCCAGGTGCTTCAGTTTAAAT 266
QY 246 GAAGTGGAGCCAAATGCAAGTTACCCCATGACTCCAGAGAACTTTTACAGCTGTGGGTTT 305
DB 267 GAAGTGGAGCCAAATGCAAGTTACCCCATGACTCCAGAGAACTTTTACAGCTGTGGGTTT 326
QY 306 ATGACAGAGATCCAGAGGGGTAGTTATCTCTGATGCGATTTTCAGGCAAAAGCTGCAGAT 365
DB 327 ATGACAGAGATCCAGAGGGGTAGTTATCTCTGATGCGATTTTCAGGCAAAAGCTGCAGAT 386
QY 366 AAATCCATTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGCAATAT 425
DB 387 AAATCCATTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGCAATAT 446
QY 426 TTACTGAAAGTGTCAATTAAGCTGTTTGGTGAAGAGTCTGCGAGCTTCGGGGAAGATAT 485
DB 447 TTACTGAAAGTGTCAATTAAGCTGTTTGGTGAAGAGTCTGCGAGCTTCGGGGAAGATAT 506
QY 486 ATTCGACTCTGTGAGAAATATTACTCTCAGAACCCCGAGGCAAGTAGACTTCTTAGAATGT 545
DB 507 ATTCGACTCTGTGAGAAATATTACTCTCAGAACCCCGAGGCAAGTAGACTTCTTAGAATGT 566
QY 546 GCAGAGAGCTAGAAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGCAAAATC 605
DB 567 GCAGAGAGCTAGAAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGCAAAATC 626

2y 606 CCAAACTTGTACCTGAAGGTTCTGTAGATGGGATACCAAGGATGGCTCTGTGATGCT 665
Db 627 CCAAACTTGTACCTGAAGGTTCTGTAGATGGGATACCAAGGATGGCTCTGTGATGCT 686
2y 666 GTCTACTTCAAGGAAAGTGGAAACTCCATTGAGAGAAACTAAATGGCTTTATCCT 725
Db 687 GTCTACTTCAAGGAAAGTGGAAACTCCATTGAGAGAAACTAAATGGCTTTATCCT 746
2y 726 TTCCTGTAAACTCGGCTCAGCGCACCTGTACAGATGATGTACTTGGTGAAGAACTA 785
Db 747 TTCCTGTAAACTCGGCTCAGCGCACCTGTACAGATGATGTACTTGGTGAAGAACTA 806
Qy 786 AACATTGATACATAGAGAACCTTAAGGCTCAGATTCAGAACTCCATATGCTGGAT 845
Db 807 AACATTGATACATAGAGAACCTTAAGGCTCAGATTCAGAACTCCATATGCTGGAT 866
Qy 846 GTTAGCATGTTCTTGTGCTTCAGATGAATGCGGATGTGCTCACTGGCTTGGAGCTG 905
Db 867 GTTAGCATGTTCTTGTGCTTCAGATGAATGCGGATGTGCTCACTGGCTTGGAGCTG 926
Qy 906 CTGGAAGTGAATTAACCTATGACAACTCAACAAGTGGACCAAGCAAAATGGCT 965
Db 927 CTGGAAGTGAATTAACCTATGACAACTCAACAAGTGGACCAAGCAAAATGGCT 986
Qy 966 GAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAGAGCAATTATGAACCTAGA 1025
Db 987 GAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAGAGCAATTATGAACCTAGA 1046
Qy 1026 TCCATTCTGAAAGCATGGCGATGGAGGAGCGCTTCAACAAGGAGCGGCAATTTCTCA 1085
Db 1047 TCCATTCTGAGAGCATGGCGATGGAGGAGCGCTTCAACAAGGAGCGGCAATTTCTCA 1106
Qy 1086 GGGATGTCGAGAGGATGACCTGTTCTTCTTGAAGTGTTCACCAAGCGCATGGTGAT 1145
Db 1107 GGGATGTCGAGAGGATGACCTGTTCTTCTTGAAGTGTTCACCAAGCGCATGGTGAT 1166
Qy 1146 GTGAATGAGAGGCGCACTGAAGCAGCGCTGGCAGAGAGGTGTTATCAGAGGAGAACT 1205
Db 1167 GTGAATGAGAGGCGCACTGAAGCAGCGCTGGCAGAGAGGTGTTATCAGAGGAGAACT 1226
Qy 1206 GGACATGAGGCGCCACAGTTGTGGCAGATCATCGTTCTTTCTTATATATGATCAAG 1265
Db 1227 GGACATGAGGCGCCACAGTTGTGGCAGATCATCGTTCTTTCTTATATGATCAAG 1286
Qy 1266 ATACCAAGTGCATTTATTTTTCGAGATTTTGTCCACCTTAAACTAAGCG 1319
Db 1287 ATACCAAGTGCATTTATTTTTCGAGATTTTTCCTCAGCTTAAACTAAGCG 1340

RESULT 14
AAQ21119 standard; cDNA; 1290 BP.

XX AC AAQ21119;
XX AC
XX AC
DT 25-MAR-2003 (revised)
DT 26-MAY-1992 (first entry)
XX PAI-2 in pDBP3.
XX
KW Plasminogen activator inhibitor; PAI; protease; uPA; receptor;
KW inhibition; urokinase; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 25..1272
CDS /*tag= a
FT /*label= PAI-2
XX
XX GB2246779-A.

12-FEB-1992.
XX 03-AUG-1990; 90GB-00017083.
XX 03-AUG-1990; 90GB-00017083.
PR (DELZ) DELTA BIOTECH LTD.
XX Balance DJ, Courtney MG;
FI WPI; 1992-051155/07.
XX P-PSDB; AAR20797.
DR Antitumour molecules for treatment of neoplasms - comprises first region
DR for binding to uPA receptor and second region for uPA inhibition.
XX Disclosure; Fig 2; 57pp; English.
XX A PAI-2 positive clone was isolated after screening a library (see
CC AAQ21117-8). A BglII site was created at the 5' end of the gene to create
CC a mutation in the third position of the second codon. A AlfrI site was
CC created at the 3' end to create mutations in the third position of the
CC last codon and in the first base of the stop codon. A clone derived with
CC this procedure was designated pDBP2. Two linkers were ligated with the
CC BglII-AfrII PAI-2 gene fragment from pDBP2 into HindIII + BamHI digested
CC m13mp19 to form pDBP3. See also AAQ21117-19 and AAQ21121-25. (Updated on
CC 25-MAR-2003 to correct PA field.)
XX SQ Sequence 1290 BP; 377 A; 289 C; 305 G; 319 T; 0 U; 0 Other;
Query Match 65.1%; Score 1242; DB 2; Length 1290;
Best Local Similarity 99.2%; Pred. No. 2.2e-247;
Matches 1248; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 58 AAACAATGGAGGATCTTGTGGCAAAACACACTCTTGGCCCTCAATTTATTCAAGCATC 117
Db 20 AAAAATGGAAGATCTTGTGGCAAAACACACTCTTGGCCCTCAATTTATTCAAGCATC 79
Qy 118 TGGCAAAAGCAAGCCCAACCCAGAAACCTCTTCTCTCCCATGAGCATCTCGTCCACA 177
Db 80 TGGCAAAAGCAAGCCCAACCCAGAAACCTCTTCTCTCCCATGAGCATCTCGTCCACA 139
Qy 178 TGGCCATGCTTACATGGGCTCCAGGGGAGCAGCAGCAAGACCATGAGGAGGCTTC 237
Db 140 TGGCCATGCTTACATGGGCTCCAGGGGAGCAGCAGCAAGACCATGAGGAGGCTTC 199
Qy 238 AGTTTAATGAAGTGGAGCCCAATGCAATGATACCCCATGACTCCAGAGNACTTACCAGCT 297
Db 200 AGTTTAATGAAGTGGAGCCCAATGCAATGATACCCCATGACTCCAGAGNACTTACCAGCT 259
Qy 298 GTGGGTTTCATGAGCAGATCCAGAAAGGTTAGTTATCTCTGATGCGATTTTGAGGCACAAG 357
Db 260 GTGGGTTTCATGAGCAGATCCAGAAAGGTTAGTTATCTCTGATGCGATTTTGAGGCACAAG 319
Qy 358 CTGCAGATAAATCCATTCATCTCGGCTCTCTCAGCTCTGCAATCAATGATCCACAG 417
Db 320 CTGCAGATAAATCCATTCATCTCGGCTCTCTCAGCTCTGCAATCAATGATCCACAG 379
Qy 418 GGAATTTATTTACTGGAAGTGTCAATAAGCTTTTGTGTGAGAAAGTCTCGAGCTCCGGG 477
Db 380 GGAATTTATTTACTGGAAGTGTCAATAAGCTTTTGTGTGAGAAAGTCTCGAGCTCCGGG 439
Qy 478 AAGAATATATTCGACTCTGTGAGAAATATTACTCTCAGAACCCAGCAGTAGACTTCC 537
Db 440 AAGAATATATTCGACTCTGTGAGAAATATTACTCTCAGAACCCAGCAGTAGACTTCC 499
Qy 538 TAGAATGTGAGAAAGCTTAGAAAAAGATTAAATTCCTGGTCAAGATCAAGCAAG 597
Db 500 TAGAATGTGAGAAAGCTTAGAAAAAGATTAAATTCCTGGTCAAGATCAAGCAAG 559
Qy 598 GCAAAATCCCAACTTGTGTACCTGAAAGTCTCTGTAGATGGGATACAGGATGGCTCTGG 657
Db 560 GCAAAATCCCAACTTGTGTACCTGAAAGTCTCTGTAGATGGGATACAGGATGGCTCTGG 619

QY 658 TGAATGCTCTCTACTTCAAAAGGAAAGTGGAATACTCCATTGTAGAGAGAACTAAATGGGC 717
Db 620 TGAATGCTCTCTACTTCAAAAGGAAAGTGGAATACTCCATTGTAGAGAGAACTAAATGGGC 679
QY 718 TTTATCCTTTCCGTGTAAATCTCGCTCAGCGCACACCTGTACAGATGATGTACTTGGTG 777
Db 680 TTTATCCTTTCCGTGTAAATCTCGCTCAGCGCACACCTGTACAGATGATGTACTTGGTG 739
QY 778 AAAAGTAAACATTTGATATCATAGAGAACTAAAGGCTCAGATTTAGAACTCCCATATG 837
Db 740 AAAAGTAAACATTTGATATCATAGAGAACTAAAGGCTCAGATTTAGAACTCCCATATG 799
QY 838 CTGGAGATGTACATGCTCTTGTGTGCTTCAGATGAAATGCGCATGTGTCCTCACTGGCT 897
Db 800 CTGGAGATGTACATGCTCTTGTGTGCTTCAGATGAAATGCGCATGTGTCCTCACTGGCT 859
QY 898 TGGAGCTGCTGGAAGTGAATTAACCTATGACAACTCAAACTGACAGTGGACCAAGAGACA 957
Db 860 TGGAGCTGCTGGAAGTGAATTAACCTATGACAACTCAAACTGACAGTGGACCAAGAGACA 919
QY 958 AAATGCTGAAGATGAAGTTGAGGTATACATACATACCCAGTTCCTCAATTAAGAGACATTATG 1017
Db 920 AATGCTGAAGATGAAGTTGAGGTATACATACATACCCAGTTCCTCAATTAAGAGACATTATG 979
QY 1018 AACTAGATCCATCTGAAAGCATGGGCATGGAGAGCGCTTCAACAGGAGCGGCCA 1077
Db 980 AACTAGATCCATCTGAAAGCATGGGCATGGAGAGCGCTTCAACAGGAGCGGCCA 1039
QY 1078 ATTCTCAGGAGTGTGGAGAGAAATGACCTGTCTTCTTCTGAAAGTGTTCACCAAGGCA 1137
Db 1040 ATTCTCAGGAGTGTGGAGAGAAATGACCTGTCTTCTTCTGAAAGTGTTCACCAAGGCA 1099
QY 1138 TGGTGAATGAATGAGAGGCGCATGAAGAGCGGCTGGACAGAGGTGTATGACAG 1197
Db 1100 TGGTGAATGAATGAGAGGCGCATGAAGAGCGGCTGGACAGAGGTGTATGACAG 1159
QY 1198 GGAGAACTGGACATGGAGGCGCCACAGTTTGTGGCAGATCATCCGTTTCTTTCTTATTA 1257
Db 1160 GGAGAACTGGACATGGAGGCGCCACAGTTTGTGGCAGATCATCCGTTTCTTTCTTATTA 1219
QY 1258 TGCATAGATAACCAAGTGCATTTTATTTTTCGCGAGATTTTGTCCACCTTAAACTA 1315
Db 1220 TGCATAGATAACCAAGTGCATTTTATTTTTCGCGAGATTTTCTTCCACCTTAAAGTGA 1277

RESULT 15
AB283797
ID AB283797 standard; cDNA; 1024 BP.
XX AC AB283797;
XX DT 14-MAY-2003 (first entry)
XX DE Toxicologically relevant human nucleotide sequence #956.
XX DE Toxicologically relevant gene; toxicological response; gene; ss.
XX KW Homo sapiens.
XX OS WO2003016500-A2.
XX PN 27-FEB-2003.
XX PD 16-AUG-2002; 2002MO-US026514.
XX PF 16-AUG-2001; 2001US-0313080P.
XX PR (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX PA Neft RE, Dunn RT, Adkins K, Pickett GG, Kier LD, Schmeiser K;
XX PI Alen P;
XX PI

DR WPI; 2003-268322/26.
XX Determining a toxicological response to an agent, useful for screening of
PT drugs, comprises comparing the expression profile of one or more human
PT toxic response genes to a reference gene expression profile indicative of
PT toxicity.
XX Claim 1; Page 298; 455pp; English.
CC The present invention describes a method (M1) for determining a
CC toxicological response to an agent, which comprises comparing the
CC expression profile of one or more human toxic response genes to a
CC reference gene expression profile indicative of toxicity, and so
CC determining the presence of a toxic response to the agent. Also
CC described: (1) an array comprising one or more polynucleotides selected
CC from the genes corresponding to the partial sequences given in AB282842
CC to AB284764, or their fragments of at least 20 nucleotides, or homologues
CC, and (2) determining if a gene putatively identified to be a toxic
CC response gene plays a role on toxic response pathways by determining the
CC expression profile of the gene after exposure of cells or a human subject
CC to a known toxic pharmaceutical or industrial agent, comprising: (a)
CC exposing cells to an agent or isolating cells from a human subject who
CC was exposed to an agent; (b) obtaining the test gene expression profile
CC for a putatively identified toxic response gene after exposure to a known
CC toxic pharmaceutical or industrial agent; and (c) comparing the test
CC profile to the expression profile of a gene with a similar function or
CC comparing the test profile to the expression profile of that gene after
CC exposure to other known toxic compounds. The methods are useful for
CC predicting and determining toxicological responses on a cellular, organ
CC or system level. The arrays comprising the human genes are useful for
CC toxicological screening of drugs, pharmaceutical compounds and chemicals
XX Sequence 1024 BP; 313 A; 233 C; 234 G; 244 T; 0 U; 0 Other;
Query Match 53.1%; Score 1014; DB 7; Length 1024;
Best Local Similarity 99.5%; Pred. No. 3, 3e-200;
Matches 1017; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 34 AATAACAGAGAAACAACAGATGAAACAATGGAGATCTTTGTGTGGCAAAACACACTCT 93
Db 3 ATTCCTCAATTTATTCAGCATCTGGCAAAAGCAAGCCCAACCCAGAACCTCTTCTCT 62
QY 94 TTCCCTCAATTTATTCAGCATCTGGCAAAAGCAAGCCCAACCCAGAACCTCTTCTCT 153
Db 63 TTCCCTCAATTTATTCAGCATCTGGCAAAAGCAAGCCCAACCCAGAACCTCTTCTCT 122
QY 154 CCCCATGGAGCATCTCTGTCCACCATGGCCATGGTCTACATGGGCTCCAGGGGAGCAGCG 213
Db 123 CCCCATGGAGCATCTCTGTCCACCATGGCCATGGTCTACATGGGCTCCAGGGGAGCAGCG 182
QY 214 AAGACCATGAGTGGCCCAAGGTGCTTCAGTTTAAATGAATGGAGCCCAATGCAGTTACCCCA 273
Db 183 AAGACCATGAGTGGCCCAAGGTGCTTCAGTTTAAATGAATGGAGCCCAATGCAGTTACCCCA 242
QY 274 TGACTCCAGAGAACTTTACAGCTGTGGGTTCATGCAGCAGATCCAGAGGGTAGTTATC 333
Db 243 TGACTCCAGAGAACTTTACAGCTGTGGGTTCATGCAGCAGATCCAGAGGGTAGTTATC 302
QY 334 CTGATGCGATTTTGCAGGCAACAGCTGCAGATAAAATCCATTTCCTTCCGCTCTCTCA 393
Db 303 CTGATGCGATTTTGCAGGCAACAGCTGCAGATAAAATCCATTTCCTTCCGCTCTCTCA 362
QY 394 GCTCTGCAATCAATGCATCCACAGGGAATTTTACTGGAAGTGTCAATAAGCTGTTG 453
Db 363 GCTCTGCAATCAATGCATCCACAGGGAATTTTACTGGAAGTGTCAATAAGCTGTTG 422
QY 454 GTGAGAAAGTCTGCGAGCTTCGGGAAGAAATATTCGACTCTGTGAGAAATATTACTCT 513
Db 423 GTGAGAAAGTCTGCGAGCTTCGGGAAGAAATATTCGACTCTGTGAGAAATATTACTCT 482
QY 514 CAGAACCCAGGAGTAGACTCTTCTAGATGTGCGAGAGAACTAGAGAAAGATTAATT 573
Db 483 CAGAACCCAGGAGTAGACTCTTCTAGATGTGCGAGAGAACTAGAGAAAGATTAATT 542

```

y 574 CCTGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACCTGTTACCTCAAGGTTCTGTAG 633
b 543 CCTGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACCTGTTACCTCAAGGTTCTGTAG 602
y 634 ATGGGATACAGGATGGTCTGGTGAATGCTGTCTACTTCAAAGGAAAGTGGAACCTC 693
b 603 ATGGGATACAGGATGGTCTGGTGAATGCTGTCTACTTCAAAGGAAAGTGGAACCTC 662
y 694 CATTTGAGAAAGAACTAAATGGGCTTTATCCCTTCCGTTGTAACCTCGGCTCAGCCACAC 753
b 663 CATTTGAGAAAGAACTAAATGGGCTTTATCCCTTCCGTTGTAACCTCGGCTCAGCCACAC 722
y 754 CTGTACAGATGATGTACTTGGTGAAGCTAAACATTGGATACATAGAAGACCTAAAGG 813
b 723 CTGTACAGATGATGTACTTGGTGAAGCTAAACATTGGATACATAGAAGACCTAAAGG 782
y 814 CTCAGATTCAGAACTCCCATATCTGGAGATGTAGCATGTTCTTGTGCTTCCAGATG 873
b 783 CTCAGATTCAGAACTCCCATATCTGGAGATGTAGCATGTTCTTGTGCTTCCAGATG 842
y 874 AAATTGCCGATGTGCCACTGGCTTGGAGCTGCTGGAAGTGAATAACCTATGACAAAC 933
b 843 AAATTGCCGATGTGCCACTGGCTTGGAGCTGCTGGAAGTGAATAACCTATGACAAAC 902
y 934 TCAACAAGTGGACCAAGACAAATGGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 993
b 903 TCAACAAGTGGACCAAGACAAATGGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 962
y 994 AGTTCAAATTAGAAGAGCATTTATGAACCTCAGATCCATTCTGAAAAGCATGGGCATGGAGG 1053
b 963 AGTTCAAATTAGAAGAGCATTTATGAACCTCAGATCCATTCTGAAAAGCATGGGCATGGAGG 1022
y 1054 AC 1055
b 1023 AC 1024

```

Search completed: March 10, 2004, 10:59:51
 Job time : 471.37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:19:42 ; Search time 87.1303 Seconds
(without alignments)
12152.440 Million cell updates/sec

Title: US-10-084-817-27

Perfect score: 1908
Sequence: 1 gagagcattgccgtcaga.....aacaaaaataaaaaagg 1908

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq*
- 5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq*
- 6: /cgn2_6/ptodata/2/ina/backfile1.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1876.6	98.4	1884	4	US-09-023-655-1188
2	1847.6	96.8	1854	4	US-09-023-655-1189
3	1847.6	96.8	2409	1	US-07-911-531-18
4	1847.6	96.8	2409	1	US-07-933-636A-18
5	1802	84.0	1810	1	US-07-968-286B-1
6	1402.6	73.5	1512	1	US-07-968-286B-3
7	1342.6	70.4	1482	1	US-07-968-286B-5
8	1249.4	65.3	1340	1	US-07-679-052A-14
9	1246	65.3	1340	1	US-07-679-052A-16
10	279	14.6	1519	4	US-10-024-427-3
11	274	14.4	1278	4	US-10-024-427-1
12	268.4	14.1	1316	1	US-08-315-831A-12
13	268.4	14.1	1316	1	US-08-662-318-12
14	268.4	14.1	1316	5	PCT-US95-12509-12
15	268.4	14.1	1430	4	US-09-016-434-710
16	266.8	14.0	1316	1	US-07-955-461A-12
17	265.8	13.9	1152	1	US-08-315-831A-13
18	265.8	13.9	1152	1	US-08-662-318-13
19	265.8	13.9	1152	5	PCT-US95-12509-13
20	252.4	13.2	1385	4	US-09-016-434-1075
21	252.4	13.2	1385	4	US-09-023-655-892
22	252.4	13.2	1393	1	US-08-464-148-3
23	252.4	13.2	1393	1	US-08-385-500-3
24	252.4	13.2	1393	1	US-08-846-784-3
25	251.6	13.2	1366	3	US-09-200-965-1
26	249.2	13.1	1361	4	US-09-016-434-1284
27	237.4	12.4	241	4	US-09-389-681-338

C 28	237.4	12.4	241	4	US-09-620-405B-338	Sequence 338, App
C 29	237.4	12.4	241	4	US-09-433-826B-338	Sequence 338, App
C 30	237.4	12.4	241	4	US-09-604-287A-338	Sequence 338, App
C 31	237.4	12.4	241	4	US-09-834-755-338	Sequence 338, App
32	234.8	12.3	1308	4	US-09-123-912-109	Sequence 109, App
33	234.8	12.3	1308	4	US-09-643-597-109	Sequence 109, App
34	234.8	12.3	1308	4	US-09-480-884A-109	Sequence 109, App
35	234.8	12.3	1308	4	US-09-542-615A-109	Sequence 109, App
36	234.8	12.3	1308	4	US-09-606-421B-109	Sequence 109, App
37	234.8	12.3	1308	4	US-09-221-107-109	Sequence 111, App
38	234.8	12.3	1419	4	US-09-123-912-111	Sequence 111, App
39	234.8	12.3	1419	4	US-09-643-597-111	Sequence 111, App
40	234.8	12.3	1419	4	US-09-480-884A-111	Sequence 111, App
41	234.8	12.3	1419	4	US-09-542-615A-111	Sequence 111, App
42	234.8	12.3	1419	4	US-09-606-421B-111	Sequence 111, App
43	234.8	12.3	1419	4	US-09-221-107-111	Sequence 111, App
44	229.8	12.0	1245	4	US-09-266-910-1	Sequence 1, Appli
45	229.4	12.0	1711	1	US-08-568-147B-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-023-655-1188
; Sequence 1188, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO. 1188:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1884 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9189544
; US-09-023-655-1188

Query Match

98.4%; Score 1876.6; DB 4; Length 1884;

Best Local Similarity 99.8%; Pred. No. 0;				Matches 1879; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
2y	9	TTGCCGCTCAGACAGCACTCAGAGAAATACCGAGAAACCAACAGATTTGAAACAAATGGAG	68				
Db	2	TTACCCGCTCAGACAGCACTCAGAGAAATACCGAGAAACCAACAGATTTGAAACAAATGGAG	61				
2y	69	GATCTTTGTTGGCAAAACACACTCTTTGGCTTCAATTTATTTCAAGATCTGGCAAAAGCA	128				
Db	62	GATCTTTGTTGGCAAAACACACTCTTTGGCTTCAATTTATTTCAAGCATCTGGCAAAAGCA	121				
2y	129	AGCCCAACCCAGAACTCTTCTCCCATGAGAGCATCTGCTCCACCATGGCCATGGTC	188				
Db	122	AGCCCAACCCAGAACTCTTCTCCCATGAGAGCATCTGCTCCACCATGGCCATGGTC	181				
Qy	189	TACATGGGCTCCAGGGGAGCAGCCGAGACCAAGATGCCCAAGGTCTTCAAGTTAAATGAA	248				
Db	182	TACATGGGCTCCAGGGGAGCAGCCGAGACCAAGATGCCCAAGGTCTTCAAGTTAAATGAA	241				
2y	249	GTGGAGCAATGCAGTTACCCCATGACTCCAGAGAACTTTTACCAGCTTGGGTTTCATG	308				
Db	242	GTGGAGCAATGCAGTTACCCCATGACTCCAGAGAACTTTTACCAGCTTGGGTTTCATG	301				
Qy	309	CAGCAGATCCAGAGGGTAGTTATCTCTGATGCGATTTTGAGGCACAAAGCTGCAGATAAA	368				
Db	302	CAGCAGATCCAGAGGGTAGTTATCTCTGATGCGATTTTGAGGCACAAAGCTGCAGATAAA	361				
Qy	369	ATCCATTATCTCTCGGCTCTCTCAGCTCTGCAATCAATCATCCACAGGAAATTTTA	428				
Db	362	ATCCATTATCTCTCGGCTCTCTCAGCTCTGCAATCAATCATCCACAGGGAATTTTA	421				
Qy	429	CTGGAAGGTGCAATAAGCTGTTTGGTGAGAGTCTGCGAGCTTCGCGGAAGATATATT	488				
Db	422	CTGGAAGGTGCAATAAGCTGTTTGGTGAGAGTCTGCGAGCTTCGCGGAAGATATATT	481				
Qy	489	CGACTCTGTCAGAAATATTACTCTCAGAACCCAGCAGTACGATTCCTAGATATGCA	548				
Db	482	CGACTCTGTCAGAAATATTACTCTCAGAACCCAGCAGTACGATTCCTAGATATGCA	541				
Qy	549	GAAAGCTAGAAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCA	608				
Db	542	GAAAGCTAGAAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCA	601				
Qy	609	ACTTTGTTACCTCAGGTTCTGATGAGGATACCAAGGATGTCCTGGTGAATGCTGTC	668				
Db	602	ACTTTGTTACCTCAGGTTCTGATGAGGATACCAAGGATGTCCTGGTGAATGCTGTC	661				
Qy	669	TACTTCAAGGAAAGTGGAAACTCCATTTGAGAGAACTTAAATGGGCTTTATCCTTTC	728				
Db	662	TACTTCAAGGAAAGTGGAAACTCCATTTGAGAGAACTTAAATGGGCTTTATCCTTTC	721				
Qy	729	CGTGTAACCTCGGCTCAGGCACTGATGATGATGATGATGATGATGATGATGATGATGAT	788				
Db	722	CGTGTAACCTCGGCTCAGGCACTGATGATGATGATGATGATGATGATGATGATGATGAT	781				
Qy	789	ATTGGATACATAGAACCTTAAGGCTCAGATTTAGAACTCCCATATGCTGGAGATGTT	848				
Db	782	ATTGGATACATAGAACCTTAAGGCTCAGATTTAGAACTCCCATATGCTGGAGATGTT	841				
Qy	849	AGCATGTTCTGTTGCTTCCAGATGAATTTCCGATGTCCTGCTGCTGCTGCTGCTGCTG	908				
Db	842	AGCATGTTCTGTTGCTTCCAGATGAATTTCCGATGTCCTGCTGCTGCTGCTGCTGCTG	901				
Qy	909	GAAAGTGAATAACCTATGACAACTCAACAAAGTGGACCAAGCAAAAGCAAAATGGCTGAA	968				
Db	902	GAAAGTGAATAACCTATGACAACTCAACAAAGTGGACCAAGCAAAAGCAAAATGGCTGAA	961				
Qy	969	GATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAGAGCATTTAGAACTCAGATCC	1028				
Db	962	GATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAGAGCATTTAGAACTCAGATCC	1021				
Qy	1029	ATTCTGAAAGCATGGGCATGGAGGAGCGCTTCAACAGGAGCGGCGCAATTTCTCAGGG	1088				

Db	1022	ATTCTGAGAGCATGGCATGGAGGAGCGCTTCAACAGAGGACGGGCAATTTCTCAGGG	1081
Qy	1089	ATGTCGGAGAGAAATGACCTGTTCTTCTGAGTGTTCACCAAGCCATGGTGGATGTG	1148
Db	1082	ATGTCGGAGAGAAATGACCTGTTCTTCTGAGTGTTCACCAAGCCATGGTGGATGTG	1141
Qy	1149	AATGAGGAGGCACTGAAGCAGCCGCTGCGACAGGAGGTGTTATGACAGGAGAACTGGA	1208
Db	1142	AATGAGGAGGCACTGAAGCAGCCGCTGCGACAGGAGGTGTTATGACAGGAGAACTGGA	1201
Qy	1209	CATGAGGCCCAACAGTTTGTGGCAGATCATCCGTTCTTTTCTTTTCTTATATGCAATGATA	1268
Db	1202	CATGAGGCCCAACAGTTTGTGGCAGATCATCCGTTCTTTTCTTTTCTTATATGCAATGATA	1261
Qy	1269	ACCAAGTGCAATTTTATTTTGTGGCAGATTTGTCTACCCCTTAAACCTAAGCGTGTCTTC	1328
Db	1262	ACCAAGTGCAATTTTATTTTGTGGCAGATTTGTCTACCCCTTAAACCTAAGCGTGTCTTC	1321
Qy	1329	TGCAAAAGATTTTGTAGATGAGCTGTGCTGCTCAGAAATGCTATTTCAAATTTGCCAAAA	1388
Db	1322	TGCAAAAGATTTTGTAGATGAGCTGTGCTGCTCAGAAATGCTATTTCAAATTTGCCAAAA	1381
Qy	1389	ATTTAGAGATTTTCTACATATTTCTGCTCTCTGAACTTCTGCTACCCACTAAT	1448
Db	1382	ATTTAGAGATTTTCTACATATTTCTGCTCTCTGAACTTCTGCTACCCACTAAT	1441
Qy	1449	AAAAACACAGAAATTAATAGACAAATTTGCTATTAATGAACATGACAACTTATTAATCATTT	1508
Db	1442	AAAAACACAGAAATTAATAGACAAATTTGCTATTAATGAACATGACAACTTATTAATCATTT	1501
Qy	1509	GGTCTTTAAATGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATTTTATA	1568
Db	1502	GGTCTTTAAATGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATTTTATA	1561
Qy	1569	ACATTAACCTTTTACTTTTGTATTTTATTTTATTAATGAGTGTGTTTAAATTAATTC	1628
Db	1562	ACATTAACCTTTTACTTTTGTATTTTATTTTATTAATGAGTGTGTTTAAATTAATTC	1621
Qy	1629	TCATGCTTAAATTAATGATGATTAATTAAGTTATAGAGCAGATGATCTGTTAAATTCCTA	1688
Db	1622	TCATGCTTAAATTAATGATGATTAATTAAGTTATAGAGCAGATGATCTGTTAAATTCCTA	1681
Qy	1689	TCTAATAAATGCTTTAAATTTGTTCTCATATGAAGATTAAGTAGTATCCCTCCATGCC	1748
Db	1682	TCTAATAAATGCTTTAAATTTGTTCTCATATGAAGATTAAGTAGTATCCCTCCATGCC	1741
Qy	1749	TTCTGTAAATTAATGCTGGAAGAAACATTAACAAATAGGCAAAATATATGTTATGTCATT	1808
Db	1742	TTCTGTAAATTAATGCTGGAAGAAACATTAACAAATAGGCAAAATATATGTTATGTCATT	1801
Qy	1809	TCTAGAAATACATACACATATATGCTGCTGATCTTATATTTCAATTCGAAGTATATAAT	1868
Db	1802	TCTAGAAATACATACACATATATGCTGCTGATCTTATATTTCAATTCGAAGTATATAAT	1861
Qy	1869	AAATTAACCTGCTTCCAAACAC	1891
Db	1862	AAATTAACCTGCTTCCAAACAC	1884

RESULT 2
US-09-023-655-1189
; Sequence 1189, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE

CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA: US/09/023,655
APPLICATION NUMBER: US/09/023,655
FILING DATE: HERewith

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1189:

SEQUENCE CHARACTERISTICS:
LENGTH: 1854 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g189546
J-023-655-1189

Query Match	96.8%;	Score 1847.6;	DB 4;	Length 1854;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1850;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
15	GTCCAGACAGCAACTCAGAGAATAACCCAGAGAACACCCAGATTGAAACAATGGAGGATCTT	74		
1	GTCCAGACAGCAACTCAGAGAATAACCCAGAGAACACCCAGATTGAAACAATGGAGGATCTT	60		
75	TGTTGTGGCAACACACTCTTTTGGCCCTCAATTTATTCAAGCATCTGCCAAAGCAAGAGCCCC	134		
61	TGTTGTGGCAACACACTCTTTTGGCCCTCAATTTATTCAAGCATCTGCCAAAGCAAGAGCCCC	120		
135	ACCCAGAAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATTGCCCATTGGTCTACATG	194		
121	ACCCAGAAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATTGCCCATTGGTCTACATG	180		
195	GGCTCCAGGGGCGACACCCGAAGACAGATGGCCAGGTGCTTCAGTTTAAATGAAGTGGGA	254		
181	GGCTCCAGGGGCGACACCCGAAGACAGATGGCCAGGTGCTTCAGTTTAAATGAAGTGGGA	240		
255	GCCAAATGCAHTTACCCCATGACTCCAGAGAACTTTACAGAGCTGTGGTTCATGACAGAG	314		
241	GCCAAATGCAHTTACCCCATGACTCCAGAGAACTTTACAGAGCTGTGGTTCATGACAGAG	300		
315	ATCCAGAAAGGTAGTTATTCCTGATGCGATTTTGACGGCCAAAGCTGCAGATAAAATCCAT	374		
301	ATCCAGAAAGGTAGTTATTCCTGATGCGATTTTGACGGCCAAAGCTGCAGATAAAATCCAT	360		
375	TCATCCTTCGGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAAATTTTACTCGGAA	434		
361	TCATCCTTCGGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAAATTTTACTCGGAA	420		
435	AGTGTCAATAAGCTGTTTGTGGTAGAAGTCTGCGAGCTTCGGGAAGAATAATATTCGACTC	494		
421	AGTGTCAATAAGCTGTTTGTGGTAGAAGTCTGCGAGCTTCGGGAAGAATAATATTCGACTC	480		
495	TGTCAGAAATATTACTCCTCAGAACCCACGGCAGTAGACTTCCTTAGAATGTGCAGAGAAG	554		

481	DB	TGTCAGAAATATTACTCTCCTCAGAAACCCAGGCAGTAGACTTCTCTAGAAATGTCGCAAGAA	540
555	QY	GCTAGAAAAAGATTAAATTCCTGGGTCAAGATCTAAAACCAAGGCCAAAATCCCAACTTG	614
541	DB	GCTAGAAAAAGATTAAATTCCTGGGTCAAGATCTAAAACCAAGGCCAAAATCCCAACTTG	600
615	QY	TTACTCTGAAGGTTCTGTAGATGGGGATACAGGATGGTCTCTGGTGAATGCTGTCTACTTC	674
601	DB	TTACTCTGAAGGTTCTGTAGATGGGGATACAGGATGGTCTCTGGTGAATGCTGTCTACTTC	660
675	QY	AAAGGAAAGTGAAGAACTCCATTTGAGAGAAACTAAATGGGCTTTATCTTTCCGTGTGA	734
661	DB	AAAGGAAAGTGAAGAACTCCATTTGAGAGAAACTAAATGGGCTTTATCTTTCCGTGTGA	720
735	QY	AACTCGGCTCAGGGCACACCTGTACAGATGATGTACTTCGGTGAAGCTAAACATTTGGA	794
721	DB	AACTCGGCTCAGGGCACACCTGTACAGATGATGTACTTCGGTGAAGCTAAACATTTGGA	780
795	QY	TACATAGAAGACCTAAAGGCTCAGATTTCTPAGAACTCCCATATGCTGGAGATGTTAGCATG	854
781	DB	TACATAGAAGACCTAAAGGCTCAGATTTCTAGAACTCCCATATGCTGGAGATGTTAGCATG	840
855	QY	TTCTTTGTTGCTTCCAGATGAATTCGGATGTGTCCACTGGCTTGAGCTGCTGGAAAGT	914
841	DB	TTCTTTGTTGCTTCCAGATGAATTTGCCATGTGTCCACTGGCTTGAGCTGCTGGAAAGT	900
915	QY	GAATAAACCATTACACAACTCAACAAGTGGACAGCAAGACAAAGACAAAATGSGCTGAAGATGA	974
901	DB	GAATAAACCATTACACAACTCAACAAGTGGACAGCAAGACAAAGATGSGCTGAAGATGA	960
975	QY	GTTGAGGTATACATACCCAGATTCCAAATPAGAAGAGCATATATGAATCAGATCCATTCTG	1034
961	DB	GTTGAGGTATACATACCCAGATTCCAAATPAGAAGAGCATATATGAATCAGATCCATTCTG	1020
1035	QY	AAAGCATGGGATGGAGAGCGCTTCAACAAGGACGGCCCAATTTCTCAGGGATGTG	1094
1021	DB	AGAGCATGGGATGGAGAGCGCTTCAACAAGGACGGCCCAATTTCTCAGGGATGTG	1080
1095	QY	GAGAGGAATGACCTGTTTCTTTCTGAAGTGTTCACCAAGCCATGGTGGATGGAATGAG	1154
1081	DB	GAGAGGAATGACCTGTTTCTTTCTGAAGTGTTCACCAAGCCATGGTGGATGGAATGAG	1140
1155	QY	GAGGCACATGAAGCAGCCCTGGCAACAGAGGTGTTATGACAGGGAGAACTGGACATGGA	1214
1141	DB	GAGGCACATGAAGCAGCCCTGGCAACAGAGGTGTTATGACAGGGAGAACTGGACATGGA	1200
1215	QY	GGCCACAGTTGTGGCAGATCATCGTTTCTTTTCTTTATATGATTAAGATAACCAAG	1274
1201	DB	GGCCACAGTTGTGGCAGATCATCGTTTCTTTTCTTTATATGATTAAGATAACCAAC	1260
1275	QY	TGCATTTTATTTTTCGGCAGATTTTGCTCACCTCAAACTAAAGCGTGTCTCTCTCAAA	1334
1261	DB	TGCATTTTATTTTTCGGCAGATTTTCTCTCACCTCAAACTAAAGCGTGTCTCTCAAA	1320
1335	QY	AGATTTTCTAGATGAGCTGTGCTCTCAGATGCTATTTCAAAATGCCAAAAATTTAG	1394
1321	DB	AGATTTTCTAGATGAGCTGTGCTCTCAGATGCTATTTCAAAATGCCAAAAATTTAG	1380
1395	QY	AGATGTTTTCTACATATTTCTGCTCTTCTTGAAACAATCTGCTACCCATPAAATAAAAAC	1454
1381	DB	AGATGTTTTCTACATATTTCTGCTCTTCTTGAAACAATCTGCTACCCATPAAATAAAAAC	1440
1455	QY	ACAGAAATAATTAGACAAATGTCTATTATAACATGACACCCCTATTAAATCATTTGGTCTT	1514
1441	DB	ACAGAAATAATTAGACAAATGTCTATTATAACATGACACCCCTATTAAATCATTTGGTCTT	1500
1515	QY	CTAAAATGGGATCATGCCCAATTAGATTTTCTTACTATCAGTTTATTTTATAACATTA	1574
1501	DB	CTAAAATGGGATCATGCCCAATTAGATTTTCTTACTATCAGTTTATTTTATAACATTA	1560
1575	QY	ACTTTTACTTTGTTATTTATTTATTTTATATAATGTTGAGTTTTTAAATATTGCTCACTG	1634
1561	DB	ACTTTTACTTTGTTATTTATTTTATATAATGTTGAGTTTTTAAATATTGCTCACTG	1620

1021	AGAAGCATGGGCATGGAGGACGCCTTCAACAAAGGACGGGCGCAATTTCTCAGGAATGTCG	1080
1095	GAGAGGAATGACCTGCTTTCTTTCTGAAGTGTTCACCAAGCCATCGTGGATGTAATGAG	1154
	GAGAGGAATGACCTGCTTTCTTTCTGAAGTGTTCACCAAGCCATCGTGGATGTAATGAG	1140
	GAGAGGAATGACCTGCTTTCTTTCTGAAGTGTTCACCAAGCCATCGTGGATGTAATGAG	1140
1155	GAGGGCACTGAAGCAGCGCTGGCACAGGAGGTGTATGACAGGGGAGAACTGGACATGGA	1214
	GAGGGCACTGAAGCAGCGCTGGCACAGGAGGTGTATGACAGGGGAGAACTGGACATGGA	1200
	GAGGGCACTGAAGCAGCGCTGGCACAGGAGGTGTATGACAGGGGAGAACTGGACATGGA	1200
1215	GGCCCAACAGTTTGTGGCAGATCATCCGTTTCTTTTTCTTATTATGCATAAGATAACCAAG	1274
	GGCCCAACAGTTTGTGGCAGATCATCCGTTTCTTTTTCTTATTATGCATAAGATAACCAAC	1260
	GGCCCAACAGTTTGTGGCAGATCATCCGTTTCTTTTTCTTATTATGCATAAGATAACCAAC	1260
1275	TGCATTTTATTTTTCGGCAGATTTTGCACACCTAAACTAGCGTGCCTCTTGCAAA	1334
	TGCATTTTATTTTTCGGCAGATTTTCTCACCCCTAAACTAGCGTGCCTCTTGCAAA	1320
	TGCATTTTATTTTTCGGCAGATTTTCTCACCCCTAAACTAGCGTGCCTCTTGCAAA	1320
1335	AGATTTTGTAGATGAGCTGTGTGCTCAGAAATGCTATTTCAAAATGGCCAAAATTTAG	1394
	AGATTTTGTAGATGAGCTGTGTGCTCAGAAATGCTATTTCAAAATGGCCAAAATTTAG	1380
	AGATTTTGTAGATGAGCTGTGTGCTCAGAAATGCTATTTCAAAATGGCCAAAATTTAG	1380
1395	AGATGTTTTCTACATATTTCTGCTCTCTGACAACTTCTGCTACCCACTAAATAAAAAAC	1454
	AGATGTTTTCTACATATTTCTGCTCTCTGACAACTTCTGCTACCCACTAAATAAAAAAC	1440
	AGATGTTTTCTACATATTTCTGCTCTCTGACAACTTCTGCTACCCACTAAATAAAAAAC	1440
1455	ACAGAAATAATTAGACAAATGCTATTATTAACATGACAAACCTATTAAATCATTTGCTCTT	1514
	ACAGAAATAATTAGACAAATGCTATTATTAACATGACAAACCTATTAAATCATTTGCTCTT	1500
	ACAGAAATAATTAGACAAATGCTATTATTAACATGACAAACCTATTAAATCATTTGCTCTT	1500
1515	CTAAATGGGATCATGCCCATTTAGATTTTCCCTATCATCGTTATTATTTTATAACATTTA	1574
	CTAAATGGGATCATGCCCATTTAGATTTTCCCTATCATCGTTATTATTTTATAACATTTA	1560
	CTAAATGGGATCATGCCCATTTAGATTTTCCCTATCATCGTTATTATTTTATAACATTTA	1560
1575	ACTTTTACTTTGTTATTTATTTATTTATATAATATGGTGAGTTTTTAAATTTATGCTCACATG	1634
	ACTTTTACTTTGTTATTTATTTATTTATATAATATGGTGAGTTTTTAAATTTATGCTCACATG	1620
	ACTTTTACTTTGTTATTTATTTATTTATATAATATGGTGAGTTTTTAAATTTATGCTCACATG	1620
1635	CCTATTTAATGTAGCTTAATAAGTTATAGAAGCAGATGATCGTTAAATTTCCCTATCTAAT	1694
	CCTATTTAATGTAGCTTAATAAGTTATAGAAGCAGATGATCGTTAAATTTCCCTATCTAAT	1680
	CCTATTTAATGTAGCTTAATAAGTTATAGAAGCAGATGATCGTTAAATTTCCCTATCTAAT	1680
1695	AAATGCCCTTTAATGTTCTCATAAAGGAAGTAAGTAGTATCCCTCCATGGCCCTTCTGT	1754
	AAATGCCCTTTAATGTTCTCATAAAGGAAGTAAGTAGTATCCCTCCATGGCCCTTCTGT	1740
	AAATGCCCTTTAATGTTCTCATAAAGGAAGTAAGTAGTATCCCTCCATGGCCCTTCTGT	1740
1755	AATAAATATCTGGAAAAACATTAACAATAGCAAAATATAGTTATGTCATTTCTAGA	1814
	AATAAATATCTGGAAAAACATTAACAATAGCAAAATATAGTTATGTCATTTCTAGA	1800
	AATAAATATCTGGAAAAACATTAACAATAGCAAAATATAGTTATGTCATTTCTAGA	1800
1815	AATACATAACACATATATGTCGTATCTTATTTCAATTTGCAAGTATATAAT	1868
	AATACATAACACATATATGTCGTATCTTATTTCAATTTGCAAGTATATAAT	1854
	AATACATAACACATATATGTCGTATCTTATTTCAATTTGCAAGTATATAAT	1854

RESULT 4
IS-07-693-636A-18

Sequence 18, Application US/07693636A
Patent No. 5426044
GENERAL INFORMATION:
APPLICANT: ANTALIS, Toni M.
APPLICANT: BARNES, Thomas M.
APPLICANT: CLARK, Michelle A.
APPLICANT: DEVINE, Peter L.
APPLICANT: GOSS, Neil H.
APPLICANT: LEHRBACH, Philip R.
TITLE OF INVENTION: RECOMBINANT PRODUCT
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESSES:
ADDRESSES: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia

COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/693,636A
FILING DATE: 19910430
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/025,815
FILING DATE: 13-MAR-1987
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/155 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2409 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA to genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 49..1296
US-07-693-636A-18

Query Match	96.8%;	Score 1847.6;	DB 1;	Length 2409;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1850;	Conservative	0;	Mismatches	4; Indels 0; Gaps 0
Qy	15	GTCCAGACAGCAACTCAGAGAATAACACAGAGAAACAACCCAGATTGCAAAACAATGGAGGATCTT	74	
Db	1	GTCCAGACAGCAACTCAGAGAATAACACAGAGAAACAACCCAGATTGAAACAATGGAGGATCTT	60	
Qy	75	TGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCC	134	
Db	61	TGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCC	120	
Qy	135	ACCAGAAACCTCTTTCCTCTCCCCATGGAGCATCTCGTCCACCATTGGCCATGGTCTACATG	194	
Db	121	ACCAGAAACCTCTTTCCTCTCCCCATGGAGCATCTCGTCCACCATTGGTCTACATG	180	
Qy	195	GGTCTCAGGGGACGACCGAGACCCAGATCGCCAAAGTGCTTCCAGTTTAAATGAAGTGGGA	254	
Db	181	GGTCTCAGGGGACGACCGAGACCCAGATCGCCAAAGTGCTTCCAGTTTAAATGAAGTGGGA	240	
Qy	255	GCCAAATGCAGTTACCCCCCATGACTCCAGAGAACTTTACCAAGCTGTGGTTCATGCAGCAG	314	
Db	241	GCCAAATGCAGTTACCCCCCATGACTCCAGAGAACTTTACCAAGCTGTGGTTCATGCAGCAG	300	
Qy	315	ATCCAGAAAGGTAGTTATCTCGATGCCATTTTGCAGGCACAAGCTGCAGATATAAATCCAT	374	
Db	301	ATCCAGAAAGGTAGTTATCTCGATGCCATTTTGCAGGCACAAGCTGCAGATATAAATCCAT	360	
Qy	375	TCATCTCTCCGCTCTCTCAGCTCTGCAAATCAATGCATCCACAGGGAATTTATTCTGGAA	434	
Db	361	TCATCTCTCCGCTCTCTCAGCTCTGCAAATCAATGCATCCACAGGGAATTTATTCTGGAA	420	
Qy	435	AGTGTCAATAAGCTGTTTGTGTGAGAAGCTGTGCGAGCTTCCGGGAAGAATATATTCCGATC	494	
Db	421	AGTGTCAATAAGCTGTTTGTGTGAGAAGCTGTGCGAGCTTCCGGGAAGAATATATTCCGATC	480	
Qy	495	TGTCAGAAATATTACTCCTCAGAAACCCAGCAGTAGACTTCTCTAGAAATGTGCAGAAAGAA	554	
Db	481	TGTCAGAAATATTACTCCTCAGAAACCCAGCAGTAGACTTCTCTAGAAATGTGCAGAAAGAA	540	

ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/157 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1512 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: Monocyte
CELL LINE: U937
IMMEDIATE SOURCE:
CLONE: BTA 1916
FEATURE:
NAME/KEY: CDS
LOCATION: 22..1200
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "Product binds to
OTHER INFORMATION: urokinase, tissue plasminogen activator"
OTHER INFORMATION: /product= "PAI-2 variant, protease sensitive site
OTHER INFORMATION: removed"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /note= "Codes for human plasminogen activator
OTHER INFORMATION: inhibitor type 2 protein in which amino acids 74
OTHER INFORMATION: to 96 inclusive have been deleted."
US-07-768-286B-3

Query Match 73.5%; Score 1402.6; DB 1; Length 1512;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 1484; Conservative 0; Mismatches 4; Indels 69; Gaps 1;

QY	62	AATGAGGATCTTTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGC	121
DB	21	AATGAGGATCTTTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGC	80
QY	122	AAAGCAAGCCCAACAGCACTCTCTCTCCATGAGCATCTCGTCCACCATGGC	181
DB	81	AAAGCAAGCCCAACAGCACTCTCTCTCCATGAGCATCTCGTCCACCATGGC	140
QY	182	CATGGTCTACATGGGTCAGGGGAGCAGCAAGAGACAGATGGCCAGGTGCTTCAGTT	241
DB	141	CATGGTCTACATGGGTCAGGGGAGCAGCAAGAGACAGATGGCCAGGTGCTTCAGTT	200
QY	242	TAATGAAGTGGAGCAATGCAGTTACCCCATGACTCCAGAGAACTTTACCAAGCTGTGG	301
DB	201	TAATGAAGTGGAGCAATGCAGTTACCCCATGACTCCA-----	240
QY	302	GTTTCAGCAGCATCAGAGGGTAGTTATCTGTGCGATTTTGCAGGCAGCTGC	361
DB	241	-----GCACAGCTGC	251
QY	362	AGATAAAATCCATTATCTCTTCAGCTCTGCAATCAATGATCCACAGGGAA	421
DB	252	AGATAAAATCCATTATCTCTTCAGCTCTGCAATCAATGATCCACAGGGAA	311
QY	422	TTATTTACTGGAAGTGCATAGCTGTTGGTGAAGTCTGGAGCTTCCGGGAAGA	481
DB	312	TTATTTACTGGAAGTGCATAGCTGTTGGTGAAGTCTGGAGCTTCCGGGAAGA	371
QY	482	ATATATTGCACTCTGTGAGAAATATTACTCTCAGAACCCAGGCAGTAGACTTCTCTAGA	541
DB	372	ATATATTGCACTCTGTGAGAAATATTACTCTCAGAACCCAGGCAGTAGACTTCTCTAGA	431

QY	542	ATGTCAGAGAGAGCTAGAAAAAGATTAAATCTCTGGTCAAGACTCAAAACCAAGGCAA	601
DB	432	ATGTCAGAGAGAGCTAGAAAAAGATTAAATCTCTGGTCAAGACTCAAAACCAAGGCAA	491
QY	602	AATCCAAACTTGTACCTGAAGTTCTGTAGATGGGATACCAAGGATGCTCTGGTCAA	661
DB	492	AATCCAAACTTGTACCTGAAGTTCTGTAGATGGGATACCAAGGATGCTCTGGTCAA	551
QY	662	TGCTGTCTACTTCAAGGAAAGTGGAAAACTCCATTTGAGAAGAACTAAATGGGCTTTA	721
DB	552	TGCTGTCTACTTCAAGGAAAGTGGAAAACTCCATTTGAGAAGAACTAAATGGGCTTTA	611
QY	722	TCCTTTCGGTGAAGACTCGGCTCAGCGCACCTGTACAGATGATGATCTTCGTGAAA	781
DB	612	TCCTTTCGGTGAAGACTCGGCTCAGCGCACCTGTACAGATGATGATCTTCGTGAAA	671
QY	782	GCTAAACATTTGGATACATAGAAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGG	841
DB	672	GCTAAACATTTGGATACATAGAAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGG	731
QY	842	AGATGTTAGCATGTTCTTGTTCCTCCAGATGAAATGCGCATGTGTCCACTGGCTTGA	901
DB	732	AGATGTTAGCATGTTCTTGTTCCTCCAGATGAAATGCGCATGTGTCCACTGGCTTGA	791
QY	902	GCTGTGAAAGTGAATAAACCCTATGACAACTCAACAGTGGACCAAGACAAAT	961
DB	792	GCTGTGAAAGTGAATAAACCCTATGACAACTCAACAGTGGACCAAGACAAAT	851
QY	962	GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGACATTTAGAACT	1021
DB	852	GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGACATTTAGAACT	911
QY	1022	CAGATCCATCTGAAAGCATGGCGATGGAGGCGCTTCAACAGGAGCGGCGCAATTT	1081
DB	912	CAGATCCATCTGAAAGCATGGCGATGGAGGCGCTTCAACAGGAGCGGCGCAATTT	971
QY	1082	CTCAGGGATGTCGGAGAGGAATGACCTGTTCTTCTGAAAGTGTTCACCAAGCCATGGT	1141
DB	972	CTCAGGGATGTCGGAGAGGAATGACCTGTTCTTCTGAAAGTGTTCACCAAGCCATGGT	1031
QY	1142	GGATGTGAATGAGGAGGCGACTGAACAGCGCGCTGGCACAGGAGGTGTATGACAGGGAG	1201
DB	1032	GGATGTGAATGAGGAGGCGACTGAACAGCGCGCTGGCACAGGAGGTGTATGACAGGGAG	1091
QY	1202	AACTGGACATGGAGGCGCCACAGTTTGTGGCAGATCATCTTTCTTTTCTTTATATGCA	1261
DB	1092	AACTGGACATGGAGGCGCCACAGTTTGTGGCAGATCATCTTTCTTTTCTTTATATGCA	1151
QY	1262	TAAGTAAACCAAGTGCATTTTATTTTCGCGAGATTTTGTCTCACCCCTAAAGCGGTG	1321
DB	1152	TAAGTAAACCAAGTGCATTTTATTTTCGCGAGATTTTGTCTCACCCCTAAAGCGGTG	1211
QY	1322	CTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTCAAAT	1381
DB	1212	CTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTCAAAT	1271
QY	1382	GCCAAAAATTAGAGATGTTTCTTACATATTTCTGCTCTTCTGAAACAACCTCTGCTACCC	1441
DB	1272	GCCAAAAATTAGAGATGTTTCTTACATATTTCTGCTCTTCTGAAACAACCTCTGCTACCC	1331
QY	1442	ACTAAATAAAACACAGAAATAATAGACAAATGCTCTATTATTAACATGACACCCCTATTA	1501
DB	1332	ACTAAATAAAACACAGAAATAATAGACAAATGCTCTATTATTAACATGACACCCCTATTA	1391
QY	1502	ATCATTTGGTCTTCTTAAAAATGGGATCATGCCAATTTAGATTTCCTTACTATCAGTTAT	1561
DB	1392	ATCATTTGGTCTTCTTAAAAATGGGATCATGCCAATTTAGATTTCCTTACTATCAGTTAT	1451
QY	1562	TTTTATAACATTTAACTTTTACTTTTATTTATTTATTTATATATATGTTGAGTTTTT	1618
DB	1452	TTTTATAACATTTAACTTTTACTTTTATTTATTTATTTATATATATGTTGAGTTTTT	1508

```

RESULT 7
IS-07-768-286B-5
Sequence 5, Application US/07768286B
Patent No. 544153
GENERAL INFORMATION:
APPLICANT: GOSS, Neil H.
APPLICANT: RICHARDSON, Michael A.
TITLE OF INVENTION: VARIANTS OF PAI-2
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07768,286B
FILING DATE: 19911011
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00603
FILING DATE: 20-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/157 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: Monocyte
CELL LINE: U937
IMMEDIATE SOURCE:
CLONE: BTA 1922
FEATURE:
NAME/KEY: CDS
LOCATION: 22...1170
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "product binds to
OTHER INFORMATION: /product= "PAI-2 variant, protease sensitive site
OTHER INFORMATION: /removed=
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /notes= "Codes for human plasminogen activator
OTHER INFORMATION: inhibitor type 2 protein in which amino acids 66
OTHER INFORMATION: to 98 inclusive have been deleted."
US-07-768-286B-5
Query Match 70.4%; Score 1342.6; DB 1; Length 1482;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1454; Conservative 0; Mismatches 4; Indels 99; Gaps 1;
Qy 62 AATGAGGATCTTTGTGGGCAACACACTCTTTTGGCCCTCAATTATTCAGGATCTGGC 121
Db 21 AATGAGGATCTTTGTGGGCAACACACTCTTTTGGCCCTCAATTATTCAGGATCTGGC 80

```

1062 AACGGACATGAGGCCACATTTTGTGCGAGATCACTCTTTCTTTTCTTTATTATGCA 1121
1262 TAAGATAACCAAGTGCATTTTATTTTTCGGCAGATTTTGTCTACCCCTAAACCTAAGCGTG 1321
1122 TAAGATAACCAAGTGCATTTTATTTTTCGGCAGATTTTGTCTACCCCTAAACCTAAGCGTG 1181
1322 CTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCGCTCAGAAATTTGCTATTCAAAAT 1381
1182 CTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCGCTCAGAAATTTGCTATTCAAAAT 1241
1382 GCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTGCTCTGAAACAACCTTGTCTACCC 1441
1242 GCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAACCTTGTCTACCC 1301
1442 ACTAAATAAAACACAGAAATTAATAGACAAATTTGCTATTATTAACATGACAAACCTTATTA 1501
1302 ACTAAATAAAACACAGAAATTAATAGACAAATTTGCTATTATTAACATGACAAACCTTATTA 1361
1502 ATCATTTGGTCTTCAAATTTGGATCATGCGCCATTTAGATTTTCTTACTATACAGTTTAT 1551
1362 ATCATTTGGTCTTCAAATTTGGATCATGCGCCATTTAGATTTTCTTACTATACAGTTTAT 1421
1562 TTTTATAACATTAACCTTTTACTTTTGTATTATTTATTTATTTATTAATTAATGCTGAGTTTT 1618
1422 TTTTATAACATTAACCTTTTACTTTTGTATTATTTATTTATTAATTAATGCTGAGTTTT 1478

RESULT 8
US-07-679-052A-14
Sequence 14, Application US/07679052A
Patent No. 5298400
GENERAL INFORMATION:
APPLICANT: WHITEFIELD, Peter L.
APPLICANT: RICHARDSON, Michael A.
APPLICANT: BURN, Clive L.
TITLE OF INVENTION: RECOMBINANT PRODUCT
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/679,052A
FILING DATE: 19910506
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/147 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1328 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid;
DESCRIPTION: Synthetic DNA oligonucleotide
FEATURE:
NAME/KEY: CDS
LOCATION: 15..1319

FEATURE:
NAME/KEY: sig_peptide
LOCATION: 15..71
OTHER INFORMATION: /product= "Synthetic signal
OTHER INFORMATION: peptide"
US-07-679-052A-14
Query Match 65.5%; Score 1249.4; DB 1; Length 1328;
Best Local Similarity 99.5%; Pred. No. 3.1e-293;
Matches 1253; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 61 CAATGGAGATCTTTGTGTGGCAAAACACACTCTTTTGGCCCTCAATTTATTTCAAGCATCTGG 120
DB 70 CTATGGAGATCTTTGTGTGGCAAAACACACTCTTTTGGCCCTCAATTTATTTCAAGCATCTGG 129
QY 121 CAAAGCAAGCCCAACCCAGAAACCTTCTTCTCTCTCCCATGGAGCATCTCGTCCACCATGG 180
DB 130 CAAAGCAAGCCCAACCCAGAAACCTTCTTCTCTCTCCCATGGAGCATCTCGTCCACCATGG 189
QY 191 CCATGGTCTACATGGGCTCCAGGGGCGACGACGAGACGAGATGGCCCAAGGTGCTTCAGT 240
DB 190 CCATGGTCTACATGGGCTCCAGGGGCGACGACGAGACGAGATGGCCCAAGGTGCTTCAGT 249
QY 241 TTAATGAAGTGGAGCCCAATGCAGTTTACCCCATGATCTCCAGAGAACTTTTACCAAGCTGTG 300
DB 250 TTAATGAAGTGGAGCCCAATGCAGTTTACCCCATGATCTCCAGAGAACTTTTACCAAGCTGTG 309
QY 301 GGTTCATGCAGCAGATCCAGBAGGGTAGTTATCTCTGATGCGATTTTTCAGGCACAAGCTG 360
DB 310 GGTTCATGCAGCAGATCCAGBAGGGTAGTTATCTCTGATGCGATTTTTCAGGCACAAGCTG 369
QY 361 CAGATAAATCCATTCATCTCTCCGCTCTCTAGCTCTGCAATCAATGATCCACAGGGA 420
DB 370 CAGATAAATCCATTCATCTCTCCGCTCTCTAGCTCTGCAATCAATGATCCACAGGGA 429
QY 421 ATTATTTACTGGAAGAGTGTCAATAAGCTGTTTGGTGAGAACTCTCGAGAGCTTCGGGGAAG 480
DB 430 ATTATTTACTGGAAGAGTGTCAATAAGCTGTTTGGTGAGAACTCTCGAGAGCTTCGGGGAAG 489
QY 481 AATATTTAGCTCTGTCTGAGAAATATTACTCTCAGAACCCCGAGGAGTAGACTTCCTAG 540
DB 490 AATATTTAGCTCTGTCTGAGAAATATTACTCTCAGAACCCCGAGGAGTAGACTTCCTAG 549
QY 541 AATGTGCAGAGAAGCTAGAAAAAGATTAAATTCCTGGGTCAAGACTCAAAACCAAGGCA 600
DB 550 AATGTGCAGAGAAGCTAGAAAAAGATTAAATTCCTGGGTCAAGACTCAAAACCAAGGCA 609
QY 601 AAATCCCAAACTTTTACCTGAAGTTCTGTAGATGGGGATACCAGGATGCTCTGGTGA 660
DB 610 AAATCCCAAACTTTTACCTGAAGTTCTGTAGATGGGGATACCAGGATGCTCTGGTGA 669
QY 661 ATGCTGTCTACTTCAAAGGAAAGTGGAAAACTCCATTGTGAGAAAGAACTAAATGGGCTTT 720
DB 670 ATGCTGTCTACTTCAAAGGAAAGTGGAAAACTCCATTGTGAGAAAGAACTAAATGGGCTTT 729
QY 721 ATCTTTTCCGTGTAATCTGGCTCAGGGACACCTGTACAGATGATGTACTTGGCTGAAA 780
DB 730 ATCTTTTCCGTGTAATCTGGCTCAGGGACACCTGTACAGATGATGTACTTGGCTGAAA 789
QY 781 AGCTAAACATTGGATACATAGAAAGACTAAAGGCTCAGATTCTAGAACTCCCATATGCTG 840
DB 790 AGCTAAACATTGGATACATAGAAAGACTAAAGGCTCAGATTCTAGAACTCCCATATGCTG 849
QY 841 GAGATGTTAGCATGTTCTTGTCTTCCAGATGAAATTTGCCGATGTGTCTCCTACTGGCTGG 900
DB 850 GAGATGTTAGCATGTTCTTGTCTTCCAGATGAAATTTGCCGATGTGTCTCCTACTGGCTGG 909
QY 901 AGCTGCTGGAAGTGAATAACCTATGACAAACTCAAAGTGGACGACGAGAAAGACAAA 960
DB 910 AGCTGCTGGAAGTGAATAACCTATGACAAACTCAAAGTGGACGACGAGAAAGACAAA 969
QY 961 TGGCTGAGATGAAGTTGAGGTATACATCCCGAGTTCAATTTAGAAAGAGATTTATGAAC 1020

970 TGGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAATTAGAAGACATTATGAAC 1029
 1021 TCAGATCCATTCTGAAAGCATGGGCATGGAGAGCGCTTCAACAGGAGCGGCCAATT 1080
 1030 TCAGATCCATTCTGAAAGCATGGGCATGGAGAGCGCTTCAACAGGAGCGGCCAATT 1089
 1081 TCTCAGGATGTCGGAGAGGAATGACCTGTTTCTTTCTGAAGTGTCCACCAAGCCATGG 1140
 1090 TCTCAGGATGTCGGAGAGGAATGACCTGTTTCTTTCTGAAGTGTCCACCAAGCCATGG 1149
 1141 TGGATGTGAATGAGGAGGCGCTGAAGCAGCGCTGGGCAGAGGTGTATGACAGGGA 1200
 1150 TGGATGTGAATGAGGAGGCGCTGAAGCAGCGCTGGGCAGAGGTGTATGACAGGGA 1209
 1201 GAACTGGACATGGAGGCGCCACAGTTTGTGGCAGATCATCCGTTTCTTTTCTTATTATGC 1260
 1210 GAACTGGACATGGAGGCGCCACAGTTTGTGGCAGATCATCCGTTTCTTTTCTTATTATGC 1269
 1261 ATAGATAACCAAGTGCATTTATTTTTCGGCAGATTTTGCCTCACCTAAACCTAAGCG 1319
 1270 ATAGATAACCAAGTGCATTTATTTTTCGGCAGATTTTTCCTCACCTAAACCTAAGCG 1328

RESULT 9
 JS-07-679-052A-16
 : Sequence 16, Application US/07679052A
 : Patent No. 5298400
 : GENERAL INFORMATION:
 : APPLICANT: WHITEFIELD, Peter L.
 : APPLICANT: RICHARDSON, Michael A.
 : APPLICANT: BUNN, Clive L.
 : TITLE OF INVENTION: RECOMBINANT PRODUCT
 : NUMBER OF SEQUENCES: 17
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Foley & Lardner
 : STREET: 1800 Diagonal Road, Suite 500
 : CITY: Alexandria
 : STATE: Virginia
 : COUNTRY: USA
 : ZIP: 22313-0299
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent in Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/07/679,052A
 : FILING DATE: 19910506
 : CLASSIFICATION: 514
 : ATTORNEY/AGENT INFORMATION:
 : NAME: BENT, Stephen A.
 : REGISTRATION NUMBER: 29,768
 : REFERENCE/DOCKET NUMBER: 16786/147 CHAC
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (703) 836-9300
 : TELEFAX: (703) 683-4109
 : TELEX: 899149
 : INFORMATION FOR SEQ ID NO: 16:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1340 base pairs
 : TYPE: NUCLEIC ACID
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : MOLECULE TYPE: Other nucleic acid;
 : DESCRIPTION: Synthetic DNA oligonucleotide
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: 15...1331
 : FEATURE:
 : NAME/KEY: sig_peptide
 : LOCATION: 15...86
 : OTHER INFORMATION: /product= "a-1-antitrypsin signal"
 US-07-679-052A-16

Query Match 65.3%; Score 1246; DB 1; Length 1340;
 Best Local Similarity 99.6%; Pred. No. 28-292; 5; Indels 0; Gaps 0;
 Matches 1249; Conservative 0; Mismatches 0;
 Qy 66 GAGGATCTTTGTGGCAAAACACACTCTTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAA 125
 Db 87 GAAGATCTTTGTGGCAAAACACACTCTTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAA 146
 Qy 126 GCAAGCCCCACCCAGAACCTCTTCTCTCCCATGAGCATCTCGTCCACCATGGCCATG 185
 Db 147 GCAAGCCCCACCCAGAACCTCTTCTCTCCCATGAGCATCTCGTCCACCATGGCCATG 206
 Qy 186 GTCTACATGGGTCACAGGGGAGACCCAGAACCCAGATGSCCAAGGTGCTTCAGTTTAA 245
 Db 207 GTCTACATGGGTCACAGGGGAGACCCAGAACCCAGATGSCCAAGGTGCTTCAGTTTAA 266
 Qy 246 GAAGTGGGAGCCAAATGCAAGTACCCCATGACTCCAGAGAACTTTACAGCTGTGGGTTTC 305
 Db 267 GAAGTGGGAGCCAAATGCAAGTACCCCATGACTCCAGAGAACTTTACAGCTGTGGGTTTC 326
 Qy 306 ATGACGAGATCCAGAAAGGTAGTTATCTCTGATGCGATTTTGACGACACAGCTGCAGAT 365
 Db 327 ATGACGAGATCCAGAAAGGTAGTTATCTCTGATGCGATTTTGACGACACAGCTGCAGAT 386
 Qy 366 AAAATCCATTCATCTCTCGCTCTCTCAGCTCTGCAATCAATCAATCAATCAATCAAT 425
 Db 387 AAAATCCATTCATCTCTCGCTCTCTCAGCTCTGCAATCAATCAATCAATCAATCAAT 446
 Qy 426 TTACTGAAAGTGTCAATAAGCTGTTTGTGAGAGTCTCGAGACTTCCGGGAGAAATAT 485
 Db 447 TTACTGAAAGTGTCAATAAGCTGTTTGTGAGAGTCTCGAGACTTCCGGGAGAAATAT 506
 Qy 486 ATTGACTCTGTGAGAAATATTACTCTCTCAGAACCCAGGAGTGTCTCTGAGATGT 545
 Db 507 ATTGACTCTGTGAGAAATATTACTCTCTCAGAACCCAGGAGTGTCTCTGAGATGT 566
 Qy 546 GCAGAAAGCTAGAAAAAGATTAATCTCTGGTCAAGACTCAAGCTCAAAAGGAGGAAATC 605
 Db 567 GCAGAAAGCTAGAAAAAGATTAATCTCTGGTCAAGACTCAAGCTCAAAAGGAGGAAATC 626
 Qy 606 CCAAACTTGTACCTGAAGGTTTCTGTAGATGGGATACCAAGGATGGTCTTGGTGAATGCT 665
 Db 627 CCAAACTTGTACCTGAAGGTTTCTGTAGATGGGATACCAAGGATGGTCTTGGTGAATGCT 686
 Qy 666 GTCTACTTCAAGGAAAGTGGAAACTCCATTTGAGAGAACTTAATGGGCTTTATCCT 725
 Db 687 GTCTACTTCAAGGAAAGTGGAAACTCCATTTGAGAGAACTTAATGGGCTTTATCCT 746
 Qy 726 TTCCGTGTAACCTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGCCTGAAAAGCTA 785
 Db 747 TTCCGTGTAACCTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGCCTGAAAAGCTA 806
 Qy 786 AACATTGGATACATAGAGACCTAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGAT 845
 Db 807 AACATTGGATACATAGAGACCTAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGAT 866
 Qy 846 GTTAGCATGTTCTTGTGCTTCCAGATGAAATTCGCGATGTGTCCTCACTGGCTTGGAGCTG 905
 Db 867 GTTAGCATGTTCTTGTGCTTCCAGATGAAATTCGCGATGTGTCCTCACTGGCTTGGAGCTG 926
 Qy 906 CTGGAAGTGAATTAACCTATGACAACTCAACAGTGGACCCAGCAAGAACAAAATGGCT 965
 Db 927 CTGGAAGTGAATTAACCTATGACAACTCAACAGTGGACCCAGCAAGAACAAAATGGCT 986
 Qy 966 GAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTAGAAGAGCATTTATGAATCAGA 1025
 Db 987 GAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTAGAAGAGCATTTATGAATCAGA 1046
 Qy 1026 TCCATTCTGAAAAGCATGGGCATGGAGGAGCGCTTCAAGGAGGAGCGGCGCAATTTCTCA 1085
 Db 1047 TCCATTCTGAAAAGCATGGGCATGGAGGAGCGCTTCAAGGAGGAGCGGCGCAATTTCTCA 1106

QY 1086 GGGATGTCGGAGAGGAATGACCTGTTTCTTTCTGAAGTGTTCACCAAGCCATGGTGGAT 1145
 Db 1107 GGGATGTCGGAGAGGAATGACCTGTTTCTTTCTGAAGTGTTCACCAAGCCATGGTGGAT 1166
 QY 1146 GTGAATGAGGAGGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGAGAACT 1205
 Db 1167 GTGAATGAGGAGGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGAGAACT 1226
 QY 1206 GGACATGGAGGCCACAGTTTGTGGCAGATCATCGTTTCTTTTCTTATATGATTAAG 1265
 Db 1227 GGACATGGAGGCCACAGTTTGTGGCAGATCATCGTTTCTTTTCTTATATGATTAAG 1286
 QY 1266 ATACCAAGTGCATTTATTTTGGCAGATTTTGTCCACCTAAAATAAGCG 1319
 Db 1287 ATACCAAGTGCATTTATTTTGGCAGATTTTCTCACCTAAAATAAGCG 1340

RESULT 10
 US-10-024-427-3
 ; Sequence 3, Application US/10024427
 ; Patent No. 6583269
 ; GENERAL INFORMATION:
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Nepomnichy, Boris
 ; APPLICANT: Walke, D. Wade
 ; TITLE OF INVENTION: No. 6583269el Human Protease Inhibitor and Polynucleotides Encodi
 ; CURRENT APPLICATION NUMBER: US/10/024,427
 ; FILE REFERENCE: LEX-0283-USA
 ; PRIORITY FILING DATE: 2001-12-18
 ; PRIORITY FILING DATE: 2000-12-18
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1519
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 US-10-024-427-3

Query Match 14.6%; Score 279; DB 4; Length 1519;
 Best Local Similarity 53.6%; Pred. No. 4.4e-58;
 Matches 688; Conservative 0; Mismatches 565; Indels 30; Gaps 4;

QY 60 ACATGGAGATCTTGTGTGGGAAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTG 119
 Db 142 ACAATGGACTCTCTTGTACAGCAAAACACCAATTTTGTCTTGATCTTTTCAAGAGATA 201
 QY 120 GCAAAAGCAAGCCCAACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATG 179
 Db 202 GCAAAAGATGATGTCATATAAAACATATTTTCTCTCCCTGAGCCTCTCAGCTGCCCTT 261
 QY 180 GCATGGTCTACATGGCTCCAGGGCAGCAGCAAGACCCAGATGGCCAGGCTGCTTCCAG 239
 Db 262 GGTATGGTACGCTTGGGTGTAGAGTGACAGTGACATCAGATTTGATGAGGTACTACAC 321
 QY 240 TTTAATGAAGTGGAGGCAATGCAATTA-----CCCCATGACTCCAGAGAACTTTA 291
 Db 322 TTCAACGAATTTTCCAGATGAAAGCAAGAACTTGACCTGCTGAAAGCAACAAA 381
 QY 292 CAGCTGTGGTTCATCGAGC-AGATCAGAGGGTAGTTATCTCTGATGCGATTTTGCAG 350
 Db 392 CAAAAGTGTGGTGTGACAGCTCTCTGGAGGGGCGAAGAAAACACAGAGCCTCTGGAT 441
 QY 351 GCAAGAGCTGCAGATAAAATCCAT-----TCATCCTTCGCTCTCTC 392
 Db 442 CAGCAGGCTGGTCTTTAAACAAATGAGAGCGAGCTGGTCACTGCTACTTTTGGGAGCTT 501
 QY 393 AGCTCTGCATCATGATCATCAGGGAATTTTACTTGGAAAGTGTCAATAAGCTGTTT 452
 Db 502 CTCTCMAATTAGACAGATCAAGACTGATTACACACTGAGTATTCACACAGGCTTTAT 561
 QY 453 GGTGAGAGTCTCGAGCTTCCGGGAGAAATATATTCGACTCTGTCAGAAATATTACTCC 512

Db 562 GGAGAGCAGGAATTTCCCAATCTCTCAGGATACTTAGATGGTGTGATTTCAATTTTACCAC 621
 QY 513 TCAGAACCCAGGAGTACTAGCTTCTTAGATGTGTCAGAGAGCTAGAAAAAGATTTAAT 572
 Db 622 ACAGCAATTAAGAGTGTGATTTTCCAAAAAACCCCTGAAAAATCCAGACAAGAGATTAA 681
 QY 573 TCTGGGTCAAGACTCAAAACCAAAAGCAAAATCCCAACTTTGTACCTGAAAGTTCTCTGA 632
 Db 682 TTCTGGTTGAATGTCAATCCCAAGGTAAATCAAGGAATCTTTTTCAGCAAGAGCCTATT 741
 QY 633 GATGGGATACAGGATGTCTCTGGTGAATGCTGTCTACTTCAAGGAAGTGGAAACT 692
 Db 742 AATGTCGAGACTGTGTGTGATGCTGTGTTTACTTCAAGGCCAAATGCGGAACA 801
 QY 693 CCATTTGAGAAGAACTAAATGGGCTTTATCTTTTCCGTGTAAACTCGCTCAGCGACA 752
 Db 802 TACTTTGACCATGAAAAACACGGTGGATGACCTTTCTGTCTAAATGCGAATGAAAAACAG 861
 QY 753 CCTGTACAGATGATGTCTGCGTGAAGCAATTAACATTTGGATACATAGAACCTTAAG 812
 Db 862 AGTGAAGATGATGACGCAAAAGGACTCTACAGATTTGGCTTCATAGAGGAGTGAAG 921
 QY 813 GCTCAGATCTGAACCTCCCATATGCT---GGAGATGTAGCATGTTCTTTGTGCTTCCA 869
 Db 922 GCACAGATCTGGAATGAGGTACACCAAGGGGAGGCTCAGCATGTTCTGCTGTGCCA 981
 QY 870 GATGAATTTCCCATGTTCCACTGCTGGAGCTGTGGAGTGTGGAAGTGAATAACCTATGAC 929
 Db 982 TCTCACTTAAGATTAACCTGAAGGCTCTGGAAGCTTGAAGGAATCACTATGAA 1041
 QY 930 AAATCAAGTGGAGCAAGCAAAATGCTGAAGTGAAGTGAAGTGAAGTGAAGTATACATA 989
 Db 1042 AAAATGGTGGCTGGAGCAGCTCAGAAAAACATGTGAGAAGAAATCGTGTCTCTCTTC 1101
 QY 990 CCCCAGTTCAAAATTAGAAGAGCATTAATGACTCAGATCCATTCTCAAAAGCATGGCATG 1049
 Db 1102 CCCCAGTTCAAAATTAGAAGAGCATTAATGACTCAGATCCATTCTCAAAAGCATGGCAT 1161
 QY 1050 GAGGAGCCCTTCAAGGAGCGGGCCAAATTTCTCAGGAGTGTGCGAGGAGTAAGCTG 1109
 Db 1162 ACGGATATCTTTGATGAAAGAGGGCTGATCTTACTGGAATCTCTCCAAGTCCCAATTTG 1221
 QY 1110 TTTCTTTCTGAAGTGTTCACCAAGCCATGCTGATGTAATGAGGAGGCACTGAAGCA 1169
 Db 1222 TACTTGTCAAAATTAATCCACAAACCTTTGTGGAGTGGATGAAACCGTACCCAGGCA 1281
 QY 1170 GCGCTGGCAGAGAGTGTATGACAGGGAGAACTGGAGCAATGGAGGCCCAAGTTTGTG 1229
 Db 1282 GCTGCAGCCACTGGGCTGTTGTCTCGAAAGGTCACTACGATCTTGGGTGAGTTTAAT 1341
 QY 1230 GCAGATCATCGCTTTCTTTTCTTTTATTATGATAAGATAACCAAGTGCATTTATTTTTC 1289
 Db 1342 GCAACACCCCTTTCTTTTCTTTTATTATGATAAGATAACCAAGTGCATTTATTTTTC 1341
 QY 1290 GGAGATTTTGTCTACCCCTAAA 1312
 Db 1402 GGCAGGGTCTGCTCTCTCTAAA 1424

RESULT 11
 US-10-024-427-1
 ; Sequence 1, Application US/10024427
 ; Patent No. 6583269
 ; GENERAL INFORMATION:
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Nepomnichy, Boris
 ; APPLICANT: Walke, D. Wade
 ; TITLE OF INVENTION: No. 6583269el Human Protease Inhibitor and Polynucleotides Encodi
 ; FILE REFERENCE: LEX-0283-USA
 ; CURRENT APPLICATION NUMBER: US/10/024,427
 ; PRIORITY FILING DATE: 2001-12-18
 ; PRIORITY FILING DATE: 2000-12-18

NUMBER OF SEQ ID NOS: 3
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1
 LENGTH: 1278
 TYPE: DNA
 ORGANISM: homo sapiens
 IS-10-024-427-1

Query Match 14.4%; Score 274; DB 4; Length 1278;
 Best Local Similarity 53.4%; Pred. No. 6.6e-57;
 Matches 683; Conservative 0; Mismatches 565; Indels 30; Gaps 4;

63 ATGAGGATCTTTGTGGCAACACACACTCTTTGGCCCTCAATTTATTCAGCAATCTGGCA 122
 1 ATGAGCTCTCTGTATACAGCAACACCAAAATTTGCTTTGATCTTTTCAAGAGATAGGC 60
 123 AAAGCAAGCCCAACCACTCTCTCTCCATGGAGGATCTCTCCACCATGGCC 182
 61 AAGATGATCTGTATTAACAAATATTTTCTCTCCCTGAGCCTCTCAGCTGCCCTTGGT 120
 183 ATGCTCTACATGGGCTCCAGGGGAGCAGCCAGACAGATGGCCAAAGTGTCTTCATGTTT 242
 121 ATGCTAGCTCTGGGTGTAGAAAGTGCACATGATGATGATGATGATGATGATGATGATG 180
 243 AATGAAGTGGAGCAATGCAAGTTA-----CCCCATGATCTCCAGAGAACTTTACCA 294
 181 AACGAATTTCCAGATGAAGCAAGAACTGACCTTGTCTGAAAGCAACAA 240
 295 GCTGTGGTTCATGACAGAGA-TCCAGAGGGTATGTTATCTGATGCGATTTTGCAGGCA 353
 241 AAGTGTGCTGTGACAGCTCTCTGGAGGGGAGCAAAACACAGAGAGCTCTGGATCAG 300
 354 CAAGCTGCAGATAAATCCAT-----TCATCTTTCGCTCTCTCAGC 395
 301 CAGCTGGTCTTAAACATGAGAGCGACTGGTCACTGCTGCTTGGGAGCTTCTC 360
 396 TCTGCAATCAATGATCCACAGGGAATTTATCTGGAAGTCAATTAAGCTTTTGGT 455
 361 TCCAAATTAGACAGATCAAGACTGATTTACACACTGATGATTTGCCAACAGGCTTTATGA 420
 456 GAGAAGTCTGCGAGCTTCGGGAGCAATATATTCGACTCTGTCAGAAATATTACTCTCA 515
 421 GAGCAGGAATTTCCCAATCTGTCAAGAACTTAGATGGTGTGATTCATTTTACCACAG 480
 516 GAACCCAGGAGTATGATCTCTAGAAATGTCAGAGAGAGCTAGAAAGAAATTAATCC 575
 481 ACGATTGAAGTGTGATTTCCAAAAAACCCTGAAATAATCCAGAGAGATTAATCTTC 540
 576 TGGTCTCAGACTCAACCAAGGCAAAATCCCAACTTGTACCTGAAGTCTCTGTAGAT 635
 541 TGGTGTGATGTCAATCCAGGTAATCAAGGAATCTCTCAGCAAGGAGCTTATTAAT 600
 636 GGGATACCAAGATGCTCTGTGATGATGCTTCTTCCGTTAACTCGCTCAGCGCACACT 695
 601 GCTGAGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 696 TTTGAGAGAACTAAATGGGCTTTATCTTCCGTTAACTCGCTCAGCGCACACT 755
 661 TTTGACCATGAACACGGTGTGATGACCTTTCTGCTAAATGCGATGAACAGAGAT 720
 756 GTACAGATGATGTACTGCTGTGAAGCTTAACATTTGGATACATAGAGACCTTAAGGCT 815
 721 GTGAAGATGATGACGCAAAAGGACTCTACAGAAATTTGGCTTCATAGAGGAGGTGAAGCA 780
 816 CAGATCTTAGAATCCCATATGCT---GGAGATGTTAGCATGTTCTTGTGCTTCCAGAT 872
 781 CAGATCTTGAATGAGGTACCAAGGGAAGCTCAGCATGTTCTGCTGCTGCTGCTGCTGCT 840
 873 GAAATGCGGATGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 932
 841 CACTCTAAAGATAAAGCTGAAGGCTCTGAAGAGCTTGAAGAGGAAATCACTATGAAAAA 900
 933 CTCACAAAGTGGACGACCAAGCAAAATGGCTGAAGATGAAGTTGAGGTATACATACCC 992

Db 901 ATGTTGGCTGGAGCAGCTCAGAAAACATGTCAAGAAATCGTGTCTCTCTCTCTCTCTCT 960
 Qy 993 CAGTTCAAAATAGAAAGAGCATTTAGAACTCAGATCCATCTCTGAAAGCATGGGATGAG 1052
 Db 961 CGGTTCACTCCCTGGAAGACAGCTATGATCTCAATTCATTTTACAAGACATGGGCAATACG 1020
 Qy 1053 GACGCTTTCAACAGAGGAGCGGCCCAATTTCTCAGGATGTCCGAGAGGAATGACCTGTTT 1112
 Db 1021 GATATCTTTGATGAACAGAGGCTGATCTTACTGGAATCTCTCCAAGTCCCAATTTGTAC 1080
 Qy 1113 CTTTCTGAAGTGTTCACCAAGCCATGTTGATGTGATGTGATGTGATGTGATGTGATGTGAT 1172
 Db 1081 TTGTCAAAATTTATCCCAAAACCTTTTGTGGAGTGGATGAAACCGGTATCCCGAGCT 1140
 Qy 1173 CTTGCAAGAGAGTGTGTTATGACAGGAGAACTGAGACATGAGGAGCCCAACAGTTTGTGCA 1232
 Db 1141 GCAGCACTGGGCTGTGCTCTCGAAAGGTCACTACGATCTTTGGTGGAGTTTAAATGCC 1200
 Qy 1233 GATCATCGTTTCTTTTCTTATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1292
 Db 1201 AACCAACCTTTTCTCTTTTCAATTAAGACAAACCAAAACCAATCTCTCTTTTATGCGC 1260
 Qy 1293 AGATTTTCTCACCCCTAA 1310
 Db 1261 AGGCTCTCTCTCTTAA 1278

RESULT 12
 US-08-315-831A-12
 ; Sequence 12, Application US/08315831A
 ; Patent No. 5663299
 ; GENERAL INFORMATION:
 ; APPLICANT: Remold-O'Donnell, Eileen
 ; TITLE OF INVENTION: Human Monocyte Elastase Inhibitor
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Ave.
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/315,831A
 ; FILING DATE: September 30, 1994
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: C0279/7016
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1316 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 49..1188

OTHER INFORMATION: /codon_start= 49
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 49..1185
US-08-315-831A-12

Query Match 14.1%; Score 268.4; DB 1; Length 1316;
Best Local Similarity 55.3%; Pred. No. 1.5e-55;
Matches 543; Conservative 0; Mismatches 436; Indels 3; Gaps 1;

QY 341 GATTTTGCAGGCAAGCTGCAGATAAAATCCATTCACTCCCTCCGCTCTCTCAGCTCTGC 400
DB 216 GACTTTCCATTCAACACGGTTGAAGAGTTCAATCAAGATTCCAGAGCTTGAATGCTGA 275
QY 401 AATCAATGCATCCACAGGAAATTTTACTGGAAGTCTCAATAGCTTGTGGTGA 460
DB 276 TATCAACAACGTTGAGCGCTTATATCTGAAACTTGTCTAATAGATTATATGGAGAA 335
QY 461 GTCTCGAGCTTCCGGAAGAATATATTCGACTCTGTCAGAAATATTTACTCCTCAGAACC 520
DB 336 AACTTACAATTTCTCTCTGAGTTCTTGGTTTCGACTCAGAAACATATGGTGTGACCT 395
QY 521 CCAGGAGTAGACTTCTAGAACTGTCAGAGAAGCTAGAAAAAGATTAAATCTCTGGT 580
DB 396 GGCAGTGTGATTTTCAGCATGCTCTGAGATGCAAGAGACCAATAACCACTGGT 455
QY 581 CAAGACTCAACCAAGGCAAAATCCCAACTTGTACCTGAAGTTCTGTAGATGGGA 640
DB 456 CAAAGACACAGACAGAGAAATTCGGAACTGTGGCTTCGGGCATGGTTGATAACAT 515
QY 641 TACAGGATGGTCTGTGTAATGCTGTCTACTCTCAAGAGAAAGTGGAAACTCCATTTGA 700
DB 516 GACCAAACTGTGTAGTAATGCAATCTATTTCAAGGAAACTGGAAGGATAAATTCAT 575
QY 701 GAAGAACTAAATGGGCTTTTCTCTCCGCTGTAATCTCGCTCAGCGCACACCTGTACA 760
DB 576 GAAGAAGCCACAGCAATGCAACATTCAGATTGAAATGAAGAAGACAAAACGTGAA 635
QY 761 GATGATGATCTCGGTGAAAGCTAAACATTTGGATACATAGAACCTTAAAGGCTCAGAT 820
DB 636 AATGATGTATCAGAAGAAAAATTTGCAATATGGCTTACATCGAGGACCTTAAAGTGGCGTGT 695
QY 821 TCTAGAACTCCCATATCTCGAGATG---TTAGCATGTTCTTGTGCTTCAGATGAAT 877
DB 696 GCTGGAACTGCTTACCAAGCGAGGAGCTCAGCATGCTCTCTGCTCGCGATGACAT 755
QY 878 TGCCGATGTCTCACTGGCTTGGAGCTGCTGGAAGTGAATAACCTATGACAACTCAA 937
DB 756 TGAGGACGAGTCCACGGGCTTGAAGAAGATTGAGAACAGTTGACTTTGGAAAGTTGCA 815
QY 938 CRAAGTGACCAAGCAACAAATGGCTGAAAGATGAAGTTGAGGTATACATACCCAGTT 997
DB 816 TGAGTGGACTTAACTCAGAACTCTCGATTTTCAATGAAGTTAATGTGCTGCTGCCAGGTT 875
QY 998 CAAATTAGAGAGCATTTGAATCAGATCCATCTGAAAGCATGGGCATGGAGGACG 1057
DB 876 CAACTGGAAGAGTTTACCTCTCACTCCGCTCGACCTCGCCGCTAGTGTGACAGATCT 935
QY 1058 CTTCAACAAGGAGCGGCCAATTTCTCAGGATGTGAGAGGAATGACCTGTTCTTTTC 1117
DB 936 CTTTAAACAGTAGAAGGCTGATCTGTCTGGCATGTGAGAGCCAGAGATATTTTATATC 995
QY 1118 TGAAGTGTTCACCAAGCCATGGTGGATGTGAATGAGGAGGCACTGAAGCAGCGCGTGG 1177
DB 996 AAAAAATTGTCACAAAGTCATTTGTGGAAGTGAATGAAGAGGAAACAGAGGCGGAGCTGC 1055
QY 1178 CACAGGAGGTGTTATGACGGAGAACTGACATGAGAGGCCACAGTTTGTGGAGATCA 1237
DB 1056 CACAGGAGGATCGCAACTTTCTGCATGTTGATGCCGCCGAAGAAATTTCACTGCCGACCA 1115
QY 1238 TCCGTTTCTTTTCTTATGATGATAAGATAACCAAGTGCATTTTATTTTTCGGCAGATT 1297
DB 1116 TCCATTCCTTTCTTTATTCGGCATAAATTCCTCAGGTAGCATCTTATTTCTTGGGAGATT 1175

QY 1298 TTGCTCACCCCTAAAACCTAAGCG 1319
DB 1176 TTCTTCCCTTAGAAGAAGAG 1197

RESULT 13
US-08-662-318-12
; Sequence 12, Application US/08662318
; Patent No. 5827672
; GENERAL INFORMATION:
; APPLICANT: Remold-O'Donnell, Eileen
; TITLE OF INVENTION: Human Monocyte Elastase Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Ave.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/662,318
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/315,831
; FILING DATE: September 30, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 35,637
; REFERENCE/DOCKET NUMBER: C0279/7016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1316 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 49..1188
; OTHER INFORMATION: /codon_start= 49
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 49..1185
; US-08-662-318-12

Query Match 14.1%; Score 268.4; DB 1; Length 1316;
Best Local Similarity 55.3%; Pred. No. 1.5e-55;
Matches 543; Conservative 0; Mismatches 436; Indels 3; Gaps 1;

QY 341 GATTTTGCAGGCAAGCTGCAGATAAAATCCATTCACTCCCTCCGCTCTCTCAGCTCTGC 400
DB 216 GACTTTCCATTCAACACGGTTGAAGAGTTCAATCAAGATTCCAGAGCTTGAATGCTGA 275
QY 401 AATCAATGCATCCACAGGAAATTTTACTGGAAGTGTCAATAGCTTGTGGTGA 460
DB 276 TATCAACAACGTTGAGCGCTTATATCTGAAACTTGTCTAATAGATTATATGGAGAA 335
QY 461 GTCTCGAGCTTCCGGAAGAATATATTCGACTCTGTCAGAAATATTTACTCCTCAGAACC 520

336 AACCTACAAATTCCTTCCTGAGTCTCTGGTTTCGACTCAGAAACATATGCTGCTGACCT 395
521 CCAGCAGTAGACTTCCTAGAAATGTCAGAGAAGCTAGAAAAAGATTAATTCCTGGGT 580
396 GCGCAGTGTGATTTTCAGCATGCTCTGAAGATGCAAGGAAGACCAATAAACCAAGTGGT 455
581 CAAGACTCAAAACCAAGGCAAAATCCCAACTGTTTACCTGGAAGTTCCTGATGAGGA 640
456 CAAGGACAGACAGAGAAAGAAATTCGGNACTGTTGGCTTCGGGCATGTTGATACAT 515
641 TACCAAGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 700
516 GACCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575
701 GAAGAACTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 760
576 GAAGAGCCAGCAGCAATGCACCATTCAGATTGAATGAAGAGACAGAAACCTGTGA 635
761 GATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820
636 AATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 695
821 TCTAGAACTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 877
696 GCTGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755
878 TGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 937
756 TGAGCAGCAGTCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815
938 CAAGTGGACCAAGACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
816 TGAGTGGACTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 875
998 CAATTAAGAGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1057
876 CAATGGAAGAGAGTACACTCTCACTCCGACCTCGCCGCTGCTGCTGCTGCTGCTGCTGCT 935
1058 CTTCAACAGGACCGGCCAAATTTCTCAGGAGTGTGAGGAGGATGCTGCTGCTGCTGCTGCT 1117
936 CTTTAACTAGTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 995
1118 TGAAGTGTCCACCAAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177
996 AAAAAATGTCACAGTCACTTTGTGGAAGTGAATGAAGAGGAAACAGAGGCGGAGCTGC 1055
1178 CACAGGAGTGTATGACAGGAGAACTGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1237
1056 CACAGGAGCAGTGCACACTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1115
1238 TCCGTTCTCTTTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1297
1116 TCCATTCTCTTTCTTATTCGGCATAATTCCTCAGGTAGCATCTTATCTTGGGGAGATT 1175
1298 TTGCTCACCTAACTAAGC 1319
1176 TCTTCCCTTAGAAGAAAG 1197

RESULT 14
PCT-US95-12509-12
; Sequence 12, Application PC/TUS9512509
; GENERAL INFORMATION:
; APPLICANT: Human Monocyte Elastase Inhibitor
; TITLE OF INVENTION: Human Monocyte Elastase Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESS: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Ave.
; CITY: Boston
; STATE: MA
; COUNTRY: USA

ZIP: 02210
COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12509
FILING DATE: herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/315,831
FILING DATE: 30 September 1994
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: C0279/7016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1316 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 49..1188
OTHER INFORMATION: /codon_start= 49
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 49..1185
PCT-US95-12509-12

Query Match 14.1%; Score 268.4; DB 5; Length 1316;
Best Local Similarity 55.3%; Pred. No. 1.5e-55;
Matches 543; Conservative 0; Mismatches 436; Indels 3; Gaps 1;
QY 341 GATTTTGAGGACAGCTGACATAAATCCATTCATTCCTTCGCTCTCAGCTCTGC 400
Db 216 GACTTTCATTTCAACACGCTTCAAGAGGTTCAATTCAGATTCAGAGCTCTGAATGCTGA 275
QY 401 ATCAATGATCCACAGGGAATTTACTGGAAGTCAATAGCTGTTTGGTGAGAA 460
Db 276 TATCAACAACTGGAGCGCTTATATCTGAAACTTCTAATAGATTATATGGAGAGA 335
QY 461 GTCTGAGCTTCCGGGAAGAATATATTCGACTCTGTGAGAAATATTACTCCTCAGAAC 520
Db 336 AACTTACAATTTCTTCTGAGTCTTGTGTTTGGACTCAGAAAAACATATGCTGCTGACCT 395
QY 521 CCAGGAGTAGACTTCCCTAGAAATGTCAGAGAAGCTAGAAAAAGATTAAATTCCTGGGT 580
Db 396 GCGCAGTGTGATTTTCAGCATGCTCTGAGATGCAAGAGGACCAATAAACCAAGTGGT 455
QY 581 CAAGACTCAAAACCAAGGCAAAATCCCAACTGTTTACCTGGAAGTTCCTGATGAGGA 640
Db 456 CAAAGGACAGACAGAGGAAATTCGGAACTGTTGGCTTCGGGCATGTTGATAACAT 515
QY 641 TACCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
Db 516 GACCAAACTTGTCTAGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575
QY 701 GAAGAACTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 760
Db 576 GAAAGAGCCAGCAGCAATGCACCATTCAGATTGAATGAAGAGACAGAAACCTGTGA 635
QY 761 GATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820

```

Db 636 AATGATGATCAGAGAAAAATTTGCATATGGCTACATCGAGGACCTTAAGTCCCGTGT 695
QY 821 TCTAGAACTCCCATATGCTGGAGATG---TTAGCATGTTCTTGTGCTTCCAGATGAAT 877
Db 696 GTTGAACATGCTTACCAAGCGGAGGAGCTCAGCATGTTCTGCTGCGGATGACAT 755
QY 878 TCCGATGTTCCACTGGCTTGGAGCTGCTGGAAGTGAATAACCTATGACAAACTCAA 937
Db 756 TGAGGACGAGTCCACGGCTGAAGAAGATTGAGGAACAGTTGACTTTGGAAGAAGTTGCA 815
QY 938 CAAGTGGACCGCAAGACAAAATGCTGAAGATTGAGGATGAGTATACATACCCAGTT 997
Db 816 TGAGTGGATTAACCTGAGAAATCTCGATTTCAATTTGAAGTTAATGTCAGCTTGCCAGTT 875
QY 998 CAAATTTAGAGACATATGAATCTAGATCATTTCTGAAAGCATGGGCATGAGGAGCGC 1057
Db 876 CAAATCTGGAAGAGAGTTACACTCTCACTCCGACCTCGCCGCTAGGTGTGCGAGATCT 935
QY 1058 CTTCAACAAAGGAGCGGCGCAATTTCTCAGGATGTCGAGAGGAAATGACCTGTTCTTTTC 1117
Db 936 CTTTAACAGTAGCAAGGCTGATCTGCTGGCATGTGAGGACCGAGATATTTTATATC 995
QY 1118 TGAAGTGTTCACCAAGCCATGTTGGATGTGAATGAGAGGGGCACTGAGCAGCGCTGG 1177
Db 996 AAAAAATGTCACAAAGTCATTTTGGGAAGTGAATGAAGAGGGAACAGAGGGCGGAGCTGC 1055
QY 1178 CACAGGAGGTGTTATGACAGGGAGAACTGGACATGGAGGCGCCACAGTTTGTGGCAGATCA 1237
Db 1056 CACAGCAGCATCGCAACTTCTGTCATGTTGATGCCCGAAGAAAATTTCACTGCCGACCA 1115
QY 1238 TCGGTTCTTTTCTTATATGATAGATAAGCAAGTGCATTTTATTTTCGGCAGATT 1297
Db 1116 TCAATTCCTTTCTTATTCGGGATATTCCTCAGGTAGCATCTTATTCCTTGGGAGATT 1175
QY 1298 TTGCTCACCTTAAACTTAAGCG 1319
Db 1176 TTCTTCCCTTAGAAGAAAGAG 1197

```

RESULT 15

```

US-09-016-434-710
; Sequence 710, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Selhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166 710:
; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1430 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: EOSIHET02
; CLONE: 334703
; US-09-016-434-710

Query Match 14.1%; Score 268.4; DB 4; Length 1430;
Best Local Similarity 55.3%; Pred. No. 1.6e-55;
Matches 543; Conservative 0; Mismatches 436; Indels 3; Gaps 1;

QY 341 GATTTTTCGAGGCAACAGCTGCAGATAAAATCCATTTTCCTTCCTCTCTCAGCTCTGC 400
Db 348 GACTTTTCCATTTTCAACACGGTTTGAAGAGGTTTCATTTCAAGATTCCAGAGCTCTGAATGCTGA 407
QY 401 AATCAATGTCATCCACAGGGNAATTTTACTTGGAAAGTGTCAATAGCTGTTTGGTGAGAA 460
Db 408 TATCAACAAACGTGGAGCGCTTATATTTCTGAAACTTGTCTAATAGATTATATGAGAGAA 467
QY 461 GTCTGCGAGGTTCCGGGAAGAATATATTCGATCTGTGAGAAATATTACTCTCTCAGAAC 520
Db 468 AACTTCAAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 527
QY 521 CCAGGCGAGTAGACTTCTCTAGAAATGTGAGAGAGAGCTAGAGAAAGATTAATTTCTCTGGGT 580
Db 528 GGCAGTGTGGATTTTTCAGCATGCTCTGAGATGCAAGGAGACCATTAACCATGGTGGT 587
QY 581 CAAGACTCAAAACCAAGGCAAAATCCCAAACTTTTACTTACCTGAAAGTTCTCTGAGATGGGA 640
Db 588 CAAAGGACAGACAGAGAGGAGGAAATTTCCGGAACCTTTGGCTTCCGAGCATGGTTGATAACAT 647
QY 641 TACCAGGATGCTCTGCTGTAATGCTGTCTACTTCAAGGAAAGTGGAAACATCCCATTTGA 700
Db 648 GACCAAACTGTGCTAGTAATGCCATCTATTTCAAGGAAACATGGAAGGATTAATTCAT 707
QY 701 GAAGAACTAAATGGGCTTTTATCTTTCCTGTTAACTCGCTCAGCCACACCTGTACA 760
Db 708 GAAAGAGCCACACGCAATCCCAATTCAGATTGAATAAGAAAGACAGAAACATGTGAA 767
QY 761 GATGATGTAATTCGCTGAAAGGCTAAACATTTGGATGATACATAGAACCTTAAGGCTCAGAT 820
Db 768 AATGATGTAATTCAGAGAAAGAAATTTTGCATATGGCTACATCGAGGACCTTAAGTCCGCTGT 827
QY 821 TCTAGAACTCCCATATGCTGGAGATG---TTAGCATGTTCTTGTGCTTCCAGATGAAT 877
Db 828 GCTGGAACATGCTTTACCAAGCGAGGAGCTCAGCATGTTCTCTCTGCTGCGGATGACAT 887
QY 878 TGCCGATGTTCCACTGGCTTGGAGCTGCTGGAAGTGAATAACCTATGACAAACTCAA 937
Db 888 TGAGGACGAGTCCACGGGCTTGAAGAAGATTGAGGAACAGTTGACTTTTGGAAAGATTTGCA 947
QY 938 CAAAGTGGACCAAGCAACAAATGCGCTGAAGATGAACTTGAAGTATACATACCCAGTT 997
Db 948 TGAGTGGACTTAACCTGAGAAATCTCGATTTTCAATTTGAAGTTAATGTCAGCTTGCCAGTT 1007
QY 998 CAAATTTAGAGACATTTGAATCTCAGATCCATTTCTGAAAGCATGGGCATGAGAGACGC 1057
Db 1008 CAAACTGGAAGAGAGTTTACATCTCTCAACTCCGACCTCCCGCCCTAGGTGTGAGATCT 1067
QY 1058 CTTCAACAAAGGAGCGGCGCAATTTCTCAGGATGTCGAGAGGAAATGACCTGTTCTTTTC 1117
Db 1068 CTTTAAACAGTAGCAAGGCTGATCTGTCTGGCATGTGAGGACCGAGATATTTTATATC 1127
QY 1118 TGAAGTGTTCACCAAGCCATGTTGGATGTGAATGAGAGGCGCTGAGAGCGCGCTGC 1177
Db 1128 AAAAAATTTGTCACAAAGTCATTTTGTGGAAGTGAATGAAGAGGGAACAGAGCGCGCAGCTGC 1187

```

```

2y 1178 CACAGGAGGTGTTATGACAGGGAGAACTGGACATGGAGGCCCCACAGTTTGTGCGAGATCA 1237
    |||||
2b 1188 CACAGCAGGCATCGCAACTTCTGATGTTGATGCCCGAAGAAATTCACCTGCCGACCA 1247
    |||||
2y 1238 TCCGTTTCTTTTCTTATTATGACATAAGATAACCAAGTGCATTTTATTTTTCGGCGAGATT 1297
    |||||
2b 1248 TCCATTCTTTTCTTTTATTCGGCATAATTCCTCAGGTAGCATCCTATTCTTGGGGAGATT 1307
    |||||
2y 1298 TTGCTCACCCCTAAACTAAGCG 1319
    |||||
2b 1308 TTCTTCCCTTAGAAGAAAGAG 1329
    |||||

```

Search completed: March 11, 2004, 14:41:44
Job time : 93.1303 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

un on: March 11, 2004, 03:37:21; Search time 444.601 Seconds
(without alignments)
15799.545 Million cell updates/sec

itle: US-10-084-817-27

erfect score: 1908

equence: 1 gagagcattgccgcgcaga.....aacaaaaataaaaaaagg 1908

oring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

earched: 2432557 seqs, 1840798884 residues

otal number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	1908	100.0	1908	14	US-10-084-817-27
2	1897	99.4	1897	14	US-10-084-817-27
3	1896	99.4	1935	14	US-10-106-698-17
4	1872.6	98.1	1900	12	US-10-240-425-1566
5	985.8	51.7	191395	12	US-10-235-192A-45
6	984.8	51.6	1199	9	US-09-962-832-252
7	560	29.4	566	9	US-09-813-358-178
8	560	29.4	566	9	US-09-997-279-178
9	498.8	26.1	503	9	US-09-813-358-66
10	498.8	26.1	503	10	US-09-997-279-66
11	475.6	24.9	482	9	US-09-864-761-13940
12	387.6	20.3	596	14	US-10-029-386-13440
13	350.2	18.4	1256	15	US-10-027-632-122703
14	343.2	18.0	354	14	US-10-029-386-27140
15	337.2	17.7	544	14	US-10-029-386-865

16	317.8	16.7	1275	14	US-10-113-113-3	Sequence 3, Appli
17	315.4	16.5	317	9	US-09-864-761-30504	Sequence 30504, A
18	315.4	16.5	317	14	US-10-029-386-14568	Sequence 14568, A
19	276.8	14.5	3603	14	US-10-113-113-1	Sequence 1, Appli
20	274.4	14.4	314	9	US-09-813-358-186	Sequence 186, Appl
21	274.4	14.4	314	10	US-09-987-279-186	Sequence 186, Appl
22	272.4	14.3	1278	14	US-10-094-944-1	Sequence 1, Appli
23	271	14.2	1200	10	US-09-823-187-1	Sequence 1, Appli
24	269	14.1	1185	12	US-10-258-951-20	Sequence 20, Appli
25	268.4	14.1	1316	9	US-09-969-347-192	Sequence 192, Appl
26	268.4	14.1	1316	12	US-10-240-425-1339	Sequence 1339, Ap
27	268.4	14.1	1430	15	US-10-305-720-710	Sequence 710, App
28	268.4	14.1	1458	14	US-10-101-510-347	Sequence 347, App
29	268.4	14.1	1931	9	US-09-925-301-358	Sequence 358, App
30	259.2	13.6	2221	14	US-10-198-846-13963	Sequence 13963, A
31	252.4	13.2	1385	15	US-10-305-720-1075	Sequence 1075, Ap
32	252.4	13.2	1626	10	US-09-993-363-1	Sequence 1, Appli
33	250.8	13.1	1624	9	US-09-925-300-500	Sequence 500, App
34	250.8	13.1	1707	14	US-10-084-817-148	Sequence 148, App
35	249.2	13.1	1361	15	US-10-305-720-1284	Sequence 1284, Ap
36	249.2	13.1	1465	9	US-09-880-107-3027	Sequence 3027, Ap
37	240	12.6	1173	15	US-10-094-886-15	Sequence 15, Appl
38	237.4	12.4	241	9	US-09-604-287A-338	Sequence 338, App
39	237.4	12.4	241	10	US-09-551-621-338	Sequence 338, App
40	237.4	12.4	241	13	US-10-007-805-338	Sequence 338, App
41	237.4	12.4	241	14	US-10-076-622-338	Sequence 338, App
42	237.4	12.4	241	14	US-10-124-805-338	Sequence 338, App
43	234.8	12.3	1279	14	US-10-157-031-354	Sequence 354, App
44	234.8	12.3	1308	9	US-09-735-705-109	Sequence 109, App
45	234.8	12.3	1308	9	US-09-850-716A-109	Sequence 109, App

ALIGNMENTS

RESULT 1

US-10-084-817-27
; Sequence 27, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Ued G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; PRIOR FILING DATE: 2002-02-25
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 27
; LENGTH: 1908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 063646CB1
US-10-084-817-27

Query Match 100.0%; Score 1908; DB 14; Length 1908;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1908; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CAGGAGCATTGCCCGTCAGAGCAGCAACTCAGAGATAACACAGAGAACACAGATTGAAA 60
Db 1 CAGGAGCATTGCCCGTCAGAGCAGCAACTCAGAGATAACACAGAGAACACAGATTGAAA 60
Qy 61 CAATGGAGGATCTTTGTGGCAACACACTTTTCCCTCAATTTTCAAGCATCTGG 120
Db 61 CAATGGAGGATCTTTGTGGCAACACACTTTTCCCTCAATTTTCAAGCATCTGG 120

		Matches 1897; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
2y	1	GAGGAGCAATGCCCCGTCAGACAGCACTCAGAGAAATACAGAGAAACACAGATTGAAA	60		
2b	1	GAGGAGCAATGCCCCGTCAGACAGCACTCAGAGAAATACAGAGAAACACAGATTGAAA	60		
2y	61	CAATGGAGGATCTTTGTGTGGCAAAACACACTCTTTGCGCTCAATTTATTCAAGCATCTGG	120		
2b	61	CAATGGAGGATCTTTGTGTGGCAAAACACACTCTTTGCGCTCAATTTATTCAAGCATCTGG	120		
2y	121	CAAAAGGAGCCCAACCAAGCACTCTTCTCTCCCATGGAGCATCTCTCCACCATGG	180		
2b	121	CAAAAGGAGCCCAACCAAGCACTCTTCTCTCCCATGGAGCATCTCTCTCCACCATGG	180		
2y	181	CCATGGCTACATGGGCTCCAGGGGAGCAACCAAGACCCAGATGGCCAAAGTGTTCAAGT	240		
2b	181	CCATGGCTACATGGGCTCCAGGGGAGCAACCAAGACCCAGATGGCCAAAGTGTTCAAGT	240		
2y	241	TTAATGAAGTGGAGCCCAATGAGTTACCCCATGATCTCCAGAGAACTTTACCAAGCTGTG	300		
2b	241	TTAATGAAGTGGAGCCCAATGAGTTACCCCATGATCTCCAGAGAACTTTACCAAGCTGTG	300		
2y	301	GGTTCAATGAGATCCAGAAAGGAGTATCTCTGATGCGATTTTCAGGCAACAGCTG	360		
2b	301	GGTTCAATGAGATCCAGAAAGGAGTATCTCTGATGCGATTTTCAGGCAACAGCTG	360		
2y	361	CAGATAAATCCATTCATCTCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGA	420		
2b	361	CAGATAAATCCATTCATCTCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGA	420		
2y	421	ATTATTTACTGGAAGTGTCAATAAGCTGTTGGTGAGAAAGTCTGCGAGCTTCGGGAAG	480		
2b	421	ATTATTTACTGGAAGTGTCAATAAGCTGTTGGTGAGAAAGTCTGCGAGCTTCGGGAAG	480		
2y	481	AATATTTCTGCTGTGAGAAATATTAATCTCTCAGAACCCAGGAGTACGTTCTCTAG	540		
2b	481	AATATTTCTGCTGTGAGAAATATTAATCTCTCAGAACCCAGGAGTACGTTCTCTAG	540		
2y	541	AATGTGCAAGAGCTAGAAAAGATTAATCTCTGGGTGAGAACTCAACCAAGGCA	600		
2b	541	AATGTGCAAGAGCTAGAAAAGATTAATCTCTGGGTGAGAACTCAACCAAGGCA	600		
2y	601	AAATCCCAAACTTGTTACCTGGAAGTCTGTAGATGGGATACAGGATGGTCTGGTGA	660		
2b	601	AAATCCCAAACTTGTTACCTGGAAGTCTGTAGATGGGATACAGGATGGTCTGGTGA	660		
2y	661	ATGCTGTCTACTTCAAGAGAGTGAAACTCCATTTGAGAGAACTAAATGGGCTTT	720		
2b	661	ATGCTGTCTACTTCAAGAGAGTGAAACTCCATTTGAGAGAACTAAATGGGCTTT	720		
2y	721	ATCCTTTCCGTTAACTCGGCTCAGGCACACCTGTACAGATGATGTACTTGGTGAAA	780		
2b	721	ATCCTTTCCGTTAACTCGGCTCAGGCACACCTGTACAGATGATGTACTTGGTGAAA	780		
2y	781	AGCTAAACATTCGATACATAGAGACTTAAGGCTCAGATTTAGAACTCCCATATGCTG	840		
2b	781	AGCTAAACATTCGATACATAGAGACTTAAGGCTCAGATTTAGAACTCCCATATGCTG	840		
2y	841	GAGATGTTAGCATGTTCTGTTGCTTCCAGATGAAATTTGCGATGTGTCCACTGGCTGG	900		
2b	841	GAGATGTTAGCATGTTCTGTTGCTTCCAGATGAAATTTGCGATGTGTCCACTGGCTGG	900		
2y	901	AGCTCTGGAAGTGAATTAACCTATGACAACTCAACAAAGTGACCAAGCAACAA	960		
2b	901	AGCTCTGGAAGTGAATTAACCTATGACAACTCAACAAAGTGACCAAGCAACAA	960		
2y	961	TGGCTGAAGATGAAGTTGAGTATACATACCCAGTTTCAATTTAGAGAGCATTTAGAC	1020		
2b	961	TGGCTGAAGATGAAGTTGAGTATACATACCCAGTTTCAATTTAGAGAGCATTTAGAC	1020		
2y	1021	TCAGATCCATTTCTGAAAAGCATGGGATGAGAGCGCTTCAACAAAGGAGCGGCCAAT	1080		
2b	1021	TCAGATCCATTTCTGAAAAGCATGGGATGAGAGCGCTTCAACAAAGGAGCGGCCAAT	1080		

RESULT 3

US-10-106-698-261
; Sequence 261, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280

Qy	1081	TCTCAGGATGTCGAGAGGATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGG	1140
Db	1081	TCTCAGGATGTCGAGAGGATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGG	1140
Qy	1141	TGGATGTGAATGAGGAGGACACTGAAGACGCCGTGGCACAGGAGGTGTATGACAGGA	1200
Db	1141	TGGATGTGAATGAGGAGGACACTGAAGACGCCGTGGCACAGGAGGTGTATGACAGGA	1200
Qy	1201	GAACTGGACATGAGGAGGCCACACAGTTTGTGCAGATCATCCGTTCTTTCTTTATATG	1260
Db	1201	GAACTGGACATGAGGAGGCCACACAGTTTGTGCAGATCATCCGTTCTTTCTTTATATG	1260
Qy	1261	ATAAGATAACCAAGTGCAATTTATTTTTCGCGAGATTTGCTACCCCTAAACTAAGCGT	1320
Db	1261	ATAAGATAACCAAGTGCAATTTATTTTTCGCGAGATTTGCTACCCCTAAACTAAGCGT	1320
Qy	1321	GCTGCTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCCTCAGAAATGCTATTTCAAT	1380
Db	1321	GCTGCTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCCTCAGAAATGCTATTTCAAT	1380
Qy	1381	TGCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAACTTCTGCTAC	1440
Db	1381	TGCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAACTTCTGCTAC	1440
Qy	1441	CACATAAATAAACAACAGAAATTAATAGACAAATGCTATTTAATCAATGCAACCCCTAT	1500
Db	1441	CACATAAATAAACAACAGAAATTAATAGACAAATGCTATTTAATCAATGCAACCCCTAT	1500
Qy	1501	AATCATTTGGTCTCTAAATGGGATCATGCCATTTAGATTTCTTACTATGAGTTTA	1560
Db	1501	AATCATTTGGTCTCTAAATGGGATCATGCCATTTAGATTTCTTACTATGAGTTTA	1560
Qy	1561	TTTTTATACATTAACCTTTTACTTTGTTATTTATTTATTTATTAATGCTGAGTTTAA	1620
Db	1561	TTTTTATACATTAACCTTTTACTTTGTTATTTATTTATTTATTAATGCTGAGTTTAA	1620
Qy	1621	ATATGCTCAGTCCCTATTAATGAGTAAATAGTAAAGTTATAGAGCAGATGATCTGTTA	1680
Db	1621	ATATGCTCAGTCCCTATTAATGAGTAAATAGTAAAGTTATAGAGCAGATGATCTGTTA	1680
Qy	1681	ATTTCTATCTAATAAATGCTTTAATTTCTCATTAATGAAGTAAGTAGGTATCCCT	1740
Db	1681	ATTTCTATCTAATAAATGCTTTAATTTCTCATTAATGAAGTAAGTAGGTATCCCT	1740
Qy	1741	CCATGCCCTTCTGTAATAATATCTGAAAAACATTAAACAAATAGGCAATATATGTTA	1800
Db	1741	CCATGCCCTTCTGTAATAATATCTGAAAAACATTAAACAAATAGGCAATATATGTTA	1800
Qy	1801	TGTGCAATTTCTAGAAATACATAACATATATGCTGTATCTTATTTCAATTTGCAAG	1860
Db	1801	TGTGCAATTTCTAGAAATACATAACATATATGCTGTATCTTATTTCAATTTGCAAG	1860
Qy	1861	TATATAATAAATAAACCTGCTTCCAAACAAACAAAAA	1897
Db	1861	TATATAATAAATAAACCTGCTTCCAAACAAACAAAAA	1897

; PRIOR FILING DATE: 1999-11-03 ; NUMBER OF SEQ ID NOS: 8564 ; SOFTWARE: PatentIn Ver. 3.0 ; SEQ ID NO 261 ; LENGTH: 1935 ; TYPE: DNA ; ORGANISM: Homo sapiens US-10-106-698-261									
Query Match 99.4%; Score 1896; DB 14; Length 1935; Best Local Similarity 99.7%; Pred. No. 0; Matches 1899; Conservative 0; Mismatches 5; Indels 0; Gaps 0;									
QY	3	GGAGCAATGCCCTCAGACAGCAACTCAGAGAAATACAGAGAAACACAGATTGAAACA	62	903	CTGCTGGAAGTGAATAAATCTATGACAAATCTCAACAAAGTGGACGACGACAAAGCAAAATG	962	QY	903	CTGCTGGAAGTGAATAAATCTATGACAAATCTCAACAAAGTGGACGACGACAAAGCAAAATG
DB	15	GGAGCAATGCCCTCAGACAGCAACTCAGAGAAATACAGAGAAACACAGATTGAAACA	74	915	CTGCTGGAAGTGAATAAATCTATGACAAATCTCAACAAAGTGGACGACGACAAAGCAAAATG	974	DB	915	CTGCTGGAAGTGAATAAATCTATGACAAATCTCAACAAAGTGGACGACGACAAAGCAAAATG
QY	63	ATGGAGGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCA	122	963	GCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTTAGAGAGAGCAATTTGAACTC	1022	QY	963	GCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTTAGAGAGAGCAATTTGAACTC
DB	75	ATGGAGGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCA	134	975	GCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTTAGAGAGAGCAATTTGAACTC	1034	DB	975	GCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTTAGAGAGAGCAATTTGAACTC
QY	123	AAAGCAAGCCCAACCCAGCACTCTCTCCCATGGAGCATCTCGTCCACCATGGCC	182	1023	AGATCCATTTCTGAAAGCATGGGATGGAGAGCGCTTCAACAAGGGAGCGGCCCAATTTTC	1082	QY	1023	AGATCCATTTCTGAAAGCATGGGATGGAGAGCGCTTCAACAAGGGAGCGGCCCAATTTTC
DB	135	AAAGCAAGCCCAACCCAGCACTCTCTCCCATGGAGCATCTCGTCCACCATGGCC	194	1035	AGATCCATTTCTGAAAGCATGGGATGGAGAGCGCTTCAACAAGGGAGCGGCCCAATTTTC	1094	DB	1035	AGATCCATTTCTGAAAGCATGGGATGGAGAGCGCTTCAACAAGGGAGCGGCCCAATTTTC
QY	183	ATGGCTACATGGCTCCAGGGCAGCACCGAAGACCAAGATGGCAAGGTGCTTCAGTTT	242	1083	TCAGGATGTCGAGAGGAATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGGTG	1142	QY	1083	TCAGGATGTCGAGAGGAATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGGTG
DB	195	ATGGCTACATGGCTCCAGGGCAGCACCGAAGACCAAGATGGCAAGGTGCTTCAGTTT	254	1095	TCAGGATGTCGAGAGGAATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGGTG	1154	DB	1095	TCAGGATGTCGAGAGGAATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGGTG
QY	243	AATGAAGTGGAGCCCAATGAGTTACCCCATGACTCCAGAGAACTTTACCAAGCTGTGG	302	1143	GATGTGAATGAGAGGGCCTGAAAGCAGCCCTGGCAGAGGGTGTATTATGACAGGGAGA	1202	QY	1143	GATGTGAATGAGAGGGCCTGAAAGCAGCCCTGGCAGAGGGTGTATTATGACAGGGAGA
DB	255	AATGAAGTGGAGCCCAATGAGTTACCCCATGACTCCAGAGAACTTTACCAAGCTGTGG	314	1155	GATGTGAATGAGAGGGCCTGAAAGCAGCCCTGGCAGAGGGTGTATTATGACAGGGAGA	1214	DB	1155	GATGTGAATGAGAGGGCCTGAAAGCAGCCCTGGCAGAGGGTGTATTATGACAGGGAGA
QY	303	TTATCCAGAGATCCAGAGGGTGTATCTGATGCGATTTTCAGGCGCAAGCTGCA	362	1203	ACTGGACATGGAGGCCACACAGTTTGTGGCAGATCATCGTTTCTTTCTTTCTTATTTATGCA	1262	QY	1203	ACTGGACATGGAGGCCACACAGTTTGTGGCAGATCATCGTTTCTTTCTTTCTTATTTATGCA
DB	315	TTATCCAGAGATCCAGAGGGTGTATCTGATGCGATTTTCAGGCGCAAGCTGCA	374	1215	ACTGGACATGGAGGCCACACAGTTTGTGGCAGATCATCGTTTCTTTCTTTCTTATTTATGCA	1274	DB	1215	ACTGGACATGGAGGCCACACAGTTTGTGGCAGATCATCGTTTCTTTCTTTCTTATTTATGCA
QY	363	GATAAATCAATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAAT	422	1263	AGATTAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCCTAAACCTAAGCGTGC	1322	QY	1263	AGATTAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCCTAAACCTAAGCGTGC
DB	375	GATAAATCAATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAAT	434	1275	AGATTAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCCTAAACCTAAGCGTGC	1334	DB	1275	AGATTAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCCTAAACCTAAGCGTGC
QY	423	TATTTACTGGAAGTGCATTAAGCTTTTGGTGAGAGTCTGGAGCTTCGGGAAGAA	482	1323	TGCTTCTGCAAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTTCAAAATG	1382	QY	1323	TGCTTCTGCAAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTTCAAAATG
DB	435	TATTTACTGGAAGTGCATTAAGCTTTTGGTGAGAGTCTGGAGCTTCGGGAAGAA	494	1335	TGCTTCTGCAAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTTCAAAATG	1394	DB	1335	TGCTTCTGCAAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTTCAAAATG
QY	483	TATATTGCACTCTGTGAGAAATATTACTCTCTCAGAACCCCAAGGAGTAGACTTCTAGAA	542	1383	CGAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTCTGAAACAACTTCTGCTACCCA	1442	QY	1383	CGAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTCTGAAACAACTTCTGCTACCCA
DB	495	TATATTGCACTCTGTGAGAAATATTACTCTCTCAGAACCCCAAGGAGTAGACTTCTAGAA	554	1395	CGAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTCTGAAACAACTTCTGCTACCCA	1454	DB	1395	CGAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTCTGAAACAACTTCTGCTACCCA
QY	543	TGTGCAAGAGCTGAGAAAGATTAATTCCTGSGTCAAGACTCRAACCAAGGCAAA	602	1443	CTAAATTAACCAACACAGAAATAATTTAGACAAATGTCTATTATAACATGACAAACCTATTAA	1502	QY	1443	CTAAATTAACCAACACAGAAATAATTTAGACAAATGTCTATTATAACATGACAAACCTATTAA
DB	555	TGTGCAAGAGCTGAGAAAGATTAATTCCTGSGTCAAGACTCRAACCAAGGCAAA	614	1455	CTAAATTAACCAACACAGAAATAATTTAGACAAATGTCTATTATAACATGACAAACCTATTAA	1514	DB	1455	CTAAATTAACCAACACAGAAATAATTTAGACAAATGTCTATTATAACATGACAAACCTATTAA
QY	603	ATCCCAACTTGTACCTGAGGTTCTGTAGATGGGATACCAAGATGGTCTGTGTGAAT	662	1503	TCATTTGGTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATT	1562	QY	1503	TCATTTGGTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATT
DB	615	ATCCCAACTTGTACCTGAGGTTCTGTAGATGGGATACCAAGATGGTCTGTGTGAAT	674	1515	TCATTTGGTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATT	1574	DB	1515	TCATTTGGTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATT
QY	663	GCTGTCTACTTTCAAGAGAGTGGAAACTCCATTTGAGAGAACTAAATGGGCTTTAT	722	1563	TTTATAACATTAACCTTTTACTTTTATTTATTTATTTATTTATTTATTTTATTTTAAAT	1622	QY	1563	TTTATAACATTAACCTTTTACTTTTATTTATTTATTTATTTATTTTATTTTAAAT
DB	675	GCTGTCTACTTTCAAGAGAGTGGAAACTCCATTTGAGAGAACTAAATGGGCTTTAT	734	1575	TTTATAACATTAACCTTTTACTTTTATTTATTTATTTATTTATTTATTTTATTTTAAAT	1634	DB	1575	TTTATAACATTAACCTTTTACTTTTATTTATTTATTTATTTATTTATTTTATTTTAAAT
QY	723	CTTTTCGGTGTAACTCGGCTCAGGCACTGTAAGATGATGTAATTTGGTGAAG	782	1623	TATTGCTCACTGCTATTTAATGTAGCTATAAAGTTATAGAGCAGATGATCTGTTAAT	1682	QY	1623	TATTGCTCACTGCTATTTAATGTAGCTATAAAGTTATAGAGCAGATGATCTGTTAAT
DB	735	CTTTTCGGTGTAACTCGGCTCAGGCACTGTAAGATGATGTAATTTGGTGAAG	794	1635	TATTGCTCACTGCTATTTAATGTAGCTATAAAGTTATAGAGCAGATGATCTGTTAAT	1694	DB	1635	TATTGCTCACTGCTATTTAATGTAGCTATAAAGTTATAGAGCAGATGATCTGTTAAT
QY	783	CTAAACATTTGGATACATAGAGACTTAAAGGCTCAGATTTCTAGAACTCCCATATGCTGGA	842	1683	TTCCATCTTAATAATGCTTTTAAATTTGTTCTCATAATGGAAGATAGTAGGATCCCTCC	1742	QY	1683	TTCCATCTTAATAATGCTTTTAAATTTGTTCTCATAATGGAAGATAGTAGGATCCCTCC
DB	795	CTAAACATTTGGATACATAGAGACTTAAAGGCTCAGATTTCTAGAACTCCCATATGCTGGA	854	1695	TTCCATCTTAATAATGCTTTTAAATTTGTTCTCATAATGGAAGATAGTAGGATCCCTCC	1754	DB	1695	TTCCATCTTAATAATGCTTTTAAATTTGTTCTCATAATGGAAGATAGTAGGATCCCTCC
QY	843	GATGTAGCATGTTCTTTGTTCTCCAGATGAATTTGCCGATGTTCCACTGGCTTGAG	902	1743	ATGCCCTTCTGTAATAAATATCTGGAATAAAACATTAAACAATAGGCAAAATATATGTTATG	1802	QY	1743	ATGCCCTTCTGTAATAAATATCTGGAATAAAACATTAAACAATAGGCAAAATATATGTTATG
DB	855	GATGTAGCATGTTCTTTGTTCTCCAGATGAATTTGCCGATGTTCCACTGGCTTGAG	914	1755	ATGCCCTTCTGTAATAAATATCTGGAATAAAACATTAAACAATAGGCAAAATATATGTTATG	1814	DB	1755	ATGCCCTTCTGTAATAAATATCTGGAATAAAACATTAAACAATAGGCAAAATATATGTTATG
				1803	TGCATTTCTAGAAATACATTAACATATATATGCTGATCTTTATTTATTTCAATTTGCAAGTA	1862	QY	1803	TGCATTTCTAGAAATACATTAACATATATATGCTGATCTTTATTTATTTCAATTTGCAAGTA
				1815	TGCATTTCTAGAAATACATTAACATATATATGCTGATCTTTATTTATTTCAATTTGCAAGTA	1874	DB	1815	TGCATTTCTAGAAATACATTAACATATATATGCTGATCTTTATTTATTTCAATTTGCAAGTA
				1863	TATAAATAAATTAACCTGCTTCCAAACCAACAAAAATAAAAAA 1906		QY	1863	TATAAATAAATTAACCTGCTTCCAAACCAACAAAAATAAAAAA 1906
				1875	TATAAATAAATTAACCTGCTTCCAAACCAACAAAAATAAAAAA 1918		DB	1875	TATAAATAAATTAACCTGCTTCCAAACCAACAAAAATAAAAAA 1918

RESULT 4
 US-10-240-425-1566
 ; Sequence 1566, Application US/10240425
 ; Publication No. US20040033502A1

1741 CCATGCCCTTCTGTATTAATATCTGGAAGAAACATTAACAAATAGCAATATATCTTA 1800
1750 CCATGCCCTTCTGTATTAATATCTGGAAGAAACATTAACAAATAGCAATATATCTTA 1809
1801 TGTGCAATTTCTAGAAATACATACACATATATATGCTGTATCTTATATTAATTCGAAG 1860
1810 TGTGCAATTTCTAGAAATACATACACATATATATGCTGTATCTTATATTAATTCGAAG 1869
1861 TATATAAATAAATAAACCTCTCTCCAAACAAAC 1891
1870 TATATAAATAAATAAACCTCTCTCCAAACAAAC 1900

RESULT 5
US-10-235-192A-45/c
; Sequence 45, Application US/10235192A
; Publication No. US20040043389A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; TITLE OF INVENTION: Methods and Compositions for Identifying
; TITLE OF INVENTION: Risk Factors for Abnormal Lipid Levels and the Diseases
; TITLE OF INVENTION: and Disorders Associated therewith
; FILE REFERENCE: WMI-011
; CURRENT APPLICATION NUMBER: US/10/235,192A
; CURRENT FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 191395
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-235-192A-45

Query Match 51.7%; Score 985.8; DB 12; Length 191395;
Best Local Similarity 99.3%; Pred. No. 1.8e-212;
Matches 990; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 902 GCTGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGCAAGCAAAAT 961
DB 165134 GCAGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGCAAGCAAAAT 165075
QY 962 GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAAGCATTATGAAC 1021
DB 165074 GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAAGCATTATGAAC 165015
QY 1022 CAGATCCATCTGAAAAGCATGGCATGGAGGACGCTTCAACAAGGAGCGGGCAATTT 1081
DB 165014 CAGATCCATCTGAGAAGCATGGCATGGAGGACGCTTCAACAAGGAGCGGGCAATTT 164955
QY 1082 CTCAGGGATGTCGAGAGGAATGACCTGTTCTTCTGAAGTGTTCACCAAGCCATGGT 1141
DB 164954 CTCAGGGATGTCGAGAGGAATGACCTGTTCTTCTGAAGTGTTCACCAAGCCATGGT 164895
QY 1142 GGATGTGAATGAGGAGGGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGAG 1201
DB 164894 GGATGTGAATGAGGAGGGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGAG 164835
QY 1202 AACTGGACATGGAGGCCCAAGTTGTGGCAGATCATCCGTTCTTTCTTATATGCA 1261
DB 164834 AACTGGACATGGAGGCCCAAGTTGTGGCAGATCATCCCTTTCTTATATGCA 164775
QY 1262 TAAGATAACCAATGTCATTTATTTTTCGAGATTTTGTCTACCCCTAAATTAAGCGTG 1321
DB 164774 TAAGATAACCAATGTCATTTATTTTTCGAGATTTTGTCTACCCCTAAATTAAGCGTG 164715
QY 1322 CTGCTTCTGAAAAGATTTTGTAGATGAGTGTGTGCTCAGAAATTTGCTATTTCAAAAT 1381
DB 164714 CTGCTTCTGAAAAGATTTTGTAGATGAGTGTGTGCTCAGAAATTTGCTATTTCAAAAT 164655
QY 1382 GCCAAAATTTAGAGATGTTTTCTACATATTTCTGCTCTCTGAAACAACTCTGCTACCC 1441
DB 164654 GCCAAAATTTAGAGATGTTTTCTACATATTTCTGCTCTCTGAAACAACTCTGCTACCC 164595

1442 ACTAAATAAACAACAGAAATTAATTAGACAAATGTTCTATTATTAACATGACAAACCTATTA 1501
164594 ACTAAATAAACAACAGAAATTAATTAGACAAATGTTCTATTATTAACATGACAAACCTATTA 164535
1502 ATCAATTTGCTCTTCTAAATATGGGATCATGCCCAATTTAGATTTTCTTACTACTAGTTTAT 1561
164534 ATCAATTTGCTCTTCTAAATATGGGATCATGCCCAATTTAGATTTTCTTACTACTAGTTTAT 164475
1562 TTTTATAACATTAACATTTTACTTTTGTATTATTAATTTTATATATGTTGATGTTTAAAA 1621
164474 TTTTATAACATTAACATTTTACTTTTGTATTATTAATTTTATATATGTTGATGTTTAAAA 164415
1622 TTATTGCTCAGTCCCTATTTAATGTTAGCTTAATAAGTTATAGAGCAGATCATCTGTAA 1681
164414 TTATTGCTCAGTCCCTATTTAATGTTAGCTTAATAAGTTATAGAGCAGATCATCTGTAA 164355
1682 TTTCTATCTAATAAATGCTTTTAATTTGTTCTCATATGAAGATAAATAGGTATCCCTC 1741
164354 TTTCTATCTAATAAATGCTTTTAATTTGTTCTCATATGAAGATAAATAGGTATCCCTC 164295
1742 CATGCCCTTCTGTAATAAATATCTGGAAGAAACATTAACAAATAGGCAAAATATATGTTAT 1801
164294 CATGCCCTTCTGTAATAAATATCTGGAAGAAACATTAACAAATAGGCAAAATATATGTTAT 164235
1802 GTGCATTTCTAGAAATACATACATATATATGTTCTGTATCTTATATTCATTCACAAAT 1861
164234 GTGCATTTCTAGAAATACATACATATATATGTTCTGTATCTTATATTCATTCACAAAT 164175
1862 ATATAATAAATAAACCTGCTTCCAAACAAACAAAAAT 1898
164174 ATATAATAAATAAACCTGCTTCCAAACAAACAAATACAT 164138

RESULT 6
US-09-962-832-252
; Sequence 252, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-252

Query Match 51.6%; Score 984.8; DB 9; Length 1199;
Best Local Similarity 99.3%; Pred. No. 1.6e-213;
Matches 989; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 902 GCTGCTGGAAGTGAATAACCTATGCAAACTCAACAAGTGGACCAAGCAAGCAAAAT 961
DB 177 GCAGCTGGAAGTGAATAACCTATGCAAACTCAACAAGTGGACCAAGCAAGCAAAAT 236
QY 962 GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAAGAAGCATTTATGAAT 1021
DB 237 GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAAGAAGCATTTATGAAT 296
QY 1022 CAGATCCATTTCAAAAGCATGGCATTGGAGGAGCGCTTCAACAGGAGCGGGCAATTT 1081
DB 297 CAGATCCATTTCAAAAGCATGGCATTGGAGGAGCGCTTCAACAGGAGCGGGCAATTT 356
QY 1082 CTCAGGATGTCGAGAGGAATGACCTGTTCTTCTGAAAGTGTTCACCAAGCCATGGT 1141

```

b 357 CTGAGGATGTCGAGAGGAGTACCTGTTCTTCTGAAGTGTTCCCAAGCCCAATGGT 416
y 1142 GGATGTAATGAGGAGGCACTGAAGCAGCGCGTGGCAGCAGGAGTGTTATGACAGGAG 1201
b 417 GGATGTAATGAGGAGGCACTGAAGCAGCGCGTGGCAGCAGGAGTGTTATGACAGGAG 476
y 1202 AACTGGACATGGAGGCCACAGTTTGTGGCAGATCATCGTTTCTTTTCTTATTATGCA 1261
b 477 AACTGGACATGGAGGCCACAGTTTGTGGCAGATCATCGTTTCTTTTCTTATTATGCA 536
y 1262 TAAGATAACCAAGTGCAATTTATTTTTCGGCAGATTTTGTCTCACCTCAAAACCTAAGCGTG 1321
b 537 TAAGATAACCAAGTGCAATTTATTTTTCGGCAGATTTTCTCACCTCAAAACCTAAGCGTG 596
y 1322 CTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAGATTGCTATTTCAAAAT 1381
b 597 CTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAGATTGCTATTTCAAAAT 656
y 1382 GCCAAAAATTTAGAGATTTTCTACATATTTCTGCTCTTCTGAACCAACTTCTGCTACCC 1441
b 657 GCCAAAAATTTAGAGATTTTCTACATATTTCTGCTCTTCTGAACCAACTTCTGCTACCC 716
y 1442 ACTAAATATAAAACACAGAAATAATTTAGACAAATTTGCTATTATAACATGACACCCCTATTA 1501
b 717 ACTAAATATAAAACACAGAAATAATTTAGACAAATTTGCTATTATAACATGACACCCCTATTA 776
y 1502 ATCAATTTGGTCTTCAAAATGGGATCATGCCCATTTAGATTTTCTTATTAATGAGTTTAA 1561
b 777 ATCAATTTGGTCTTCAAAATGGGATCATGCCCATTTAGATTTTCTTATCAATGAGTTTAA 836
y 1562 TTTTATAACATTAATTTTACTTTTCTTATTTATTTATTTATTAATGAGTTTAA 1621
b 837 TTTTATAACATTAATTTTACTTTTCTTATTTATTTATTTATTAATGAGTTTAA 896
y 1622 TTAATGCTACTGCTTATTAATGAGTTTAAATGAGTTTAAATGAGTTTAAATGAGTTTAA 1681
b 897 TTAATGCTACTGCTTATTAATGAGTTTAAATGAGTTTAAATGAGTTTAAATGAGTTTAA 956
y 1682 TTTCTCTATCTAATAAATGCTTTTAAATGCTTTCTAATAAATGAGTTTAAATGAGTTTAA 1741
b 957 TTTCTCTATCTAATAAATGCTTTTAAATGCTTTCTAATAAATGAGTTTAAATGAGTTTAA 1016
y 1742 CATGCCCTTCTGTAATAAATCTGGAAAAACATTAACATGAGCAAAATATATGTTAT 1801
b 1017 CATGCCCTTCTGTAATAAATCTGGAAAAACATTAACATGAGCAAAATATATGTTAT 1076
y 1802 GTGCATTTCTAGAAATACATACATATATATGCTGTATCTTATTTCAATTTGCAAGT 1861
b 1077 GTGCATTTCTAGAAATACATACATATATATGCTGTATCTTATTTCAATTTGCAAGT 1136
y 1862 ATATAATAAATAACTGCTTCCCAACCAACAAAAA 1897
b 1137 ATATAATAAATAACTGCTTCCCAACCAACAAAAA 1172

```

```

; TYPE: DNA
; ORGANISM: Homo sapien
US-09-813-358-178

Query Match      29.4%; Score 560; DB 9; Length 566;
Best Local Similarity 100.0%; Pred. No. 4.1e-117;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1327 TCTCAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTCAAAATGGCAA 1386
b 566 TCTCAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTCAAAATGGCAA 507
y 1387 AAATTTAGAGATTTTCTACATATTTCTGCTCTTCTGAACAACTTCTGCTACCACTAA 1446
b 506 AAATTTAGAGATTTTCTACATATTTCTGCTCTTCTGAACAACTTCTGCTACCACTAA 447
y 1447 ATAAACACACAGAAATAATTTAGACAAATTTGCTATTATAACATGACACCCCTATTAAATCAT 1506
b 446 ATAAACACACAGAAATAATTTAGACAAATTTGCTATTATAACATGACACCCCTATTAAATCAT 387
y 1507 TTGCTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATTTTAA 1566
b 386 TTGCTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATTTTAA 327
y 1567 TAACATTAACCTTTTACTTTTCTTATTTATTTATTTATTAATGCTGAGTTTAAATTAAT 1626
b 326 TAACATTAACCTTTTACTTTTCTTATTTATTTATTTATTAATGCTGAGTTTAAATTAAT 267
y 1627 GCTCAGTCCCTTATTAATGAGTTTAAATGAGTTTAAATGAGTTTAAATGAGTTTAAAT 1686
b 266 GCTCAGTCCCTTATTAATGAGTTTAAATGAGTTTAAATGAGTTTAAATGAGTTTAAAT 207
y 1687 TATCTAATAAATGCTTTTAAATGCTTTTCTAATAAATGAGTTTAAATGAGTTTAAATGCT 1746
b 206 TATCTAATAAATGCTTTTAAATGCTTTTCTAATAAATGAGTTTAAATGAGTTTAAATGCT 147
y 1747 CCTCTGTAATAAATATCTGGAATAAATATTAACAAATGAGCAAAATATATGTTATGCA 1806
b 146 CCTCTGTAATAAATATCTGGAATAAATATTAACAAATGAGCAAAATATATGTTATGCA 87
y 1807 TTTCTAGAAATACATAACATATATATGCTGTATCTTATTTCAATTTGCAAGTATATA 1866
b 86 TTTCTAGAAATACATAACATATATATGCTGTATCTTATTTCAATTTGCAAGTATATA 27
y 1867 ATAAATAAACCTGCTTCCAA 1886
b 26 ATAAATAAACCTGCTTCCAA 7

```

```

RESULT 8
US-09-997-279-178/c
; Sequence 178, Application US/09997279
; Publication No. US20030059781A1
; GENERAL INFORMATION:
; APPLICANT: Chenault, Ruth A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.501C1
; CURRENT APPLICATION NUMBER: US/09/997,279
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 178
; LENGTH: 566
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-997-279-178

Query Match      29.4%; Score 560; DB 10; Length 566;
Best Local Similarity 100.0%; Pred. No. 4.1e-117;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```
QY 1327 TCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAGAAATTCGTATTTCAAAATGGCAA 1386
Db 566 TCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAGAAATTCGTATTTCAAAATGGCAA 507
QY 1387 AAATTTAGAGATGTTTCTACATATTTCTGCTCTCTCTGAACTTCTGCTACCCACTAA 1446
Db 506 AAATTTAGAGATGTTTCTACATATTTCTGCTCTCTCTGAACTTCTGCTACCCACTAA 447
QY 1447 ATAAAACACAGAAATAATTTAGACAATTTGCTATTATAAATGATGACCACTTAAATCAT 1506
Db 446 ATAAAACACAGAAATAATTTAGACAATTTGCTATTATAAATGATGACCACTTAAATCAT 387
QY 1507 TTGGCTCTTAAATGGGATCATGCCAATTTAGATTTTCTTACTATCAGTTTATTTTA 1566
Db 386 TTGGCTCTTAAATGGGATCATGCCAATTTAGATTTTCTTACTATCAGTTTATTTTA 327
QY 1567 TAACATTAACCTTTTACTTTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1626
Db 326 TAACATTAACCTTTTACTTTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 267
QY 1627 GCTCAGTCTTATTTATTTAGCTATAAAGTTTATAGAGCAGATGCTGTTAATTTCC 1686
Db 266 GCTCAGTCTTATTTATTTAGCTATAAAGTTTATAGAGCAGATGCTGTTAATTTCC 207
QY 1687 TATCTAATAATGCTTTAATTTGTTCTCATTAAGTAAGTAAGTATGCTTCCCTCATGC 1746
Db 206 TATCTAATAATGCTTTAATTTGTTCTCATTAAGTAAGTAAGTATGCTTCCCTCATGC 147
QY 1747 CCTCTGTAATAATCTGGAATAAATTAACATTAACATTAAGTAAGTATGCTTATGCTCA 1806
Db 146 CCTCTGTAATAATCTGGAATAAATTAACATTAACATTAAGTAAGTATGCTTATGCTCA 87
QY 1807 TTCTAGAAATACATTAACATATATATGCTGCTATCTTATTTATTTCAATGCAAGTATATA 1866
Db 86 TTCTAGAAATACATTAACATATATATGCTGCTATCTTATTTATTTCAATGCAAGTATATA 27
QY 1867 ATAAATTAACCTCTCTCCAA 1886
Db 26 ATAAATTAACCTCTCTCCAA 7

RESULT 9
US-09-813-358-66
; Sequence 66, Application US/09813358
; Patent No. US20020048759A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Pyle, Ruth
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER
; FILE REFERENCE: 210121.501
; CURRENT APPLICATION NUMBER: US/09/813.358
; NUMBER OF SEQ ID NOS: 222
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-813-358-66

Query Match 26.1%; Score 498.8; DB 9; Length 503;
Best Local Similarity 99.6%; Pred. No. 3.2e-103;
Matches 500; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 TTGCCCGTCAGACGCAACTCAGAGAATAACACAGAGAACACCAAGATTGAAACAATGGAG 68
Db 2 TCAGACCGTCAGACGCAACTCAGAGAATAACACAGAGAACACCAAGATTGAAACAATGGAG 61
QY 69 GATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCAAAAGCA 128
Db 62 GATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCAAAAGCA 121
QY 129 AGCCCAACCCAGAACCTTCTCTCTCCCATGGAGCATCTGCTCCACCATGGCCATGTC 188
Db 122 AGCCCAACCCAGAACCTTCTCTCTCCCATGGAGCATCTGCTCCACCATGGCCATGTC 181
QY 189 TACATGGGCTCCAGGGGCGAGCACCGAAGACCCAGATGGCCAAAGTGCTTCAGTTTAAATGAA 248
Db 182 TACATGGGCTCCAGGGGCGAGCACCGAAGACCCAGATGGCCAAAGTGCTTCAGTTTAAATGAA 241
QY 249 GTGGGAGCCAAATGCACTTACCCCATGACCTCCAGAGAACTTTACAGCTGTGGGTTTCATG 308
Db 242 GTGGGAGCCAAATGCACTTACCCCATGACCTCCAGAGAACTTTACAGCTGTGGGTTTCATG 301
QY 309 CAGCAGATCCAGAGGAGTAGTTATCTGATGCGATTTTGCAGGACCAAGCTGCAGATATA 368
Db 302 CAGCAGATCCAGAGGAGTAGTTATCTGATGCGATTTTGCAGGACCAAGCTGCAGATATA 361
QY 369 ATCCATTTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTATTTA 428
Db 362 ATCCATTTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTATTTA 421
QY 429 CTGGAAAGTGTCAATAGCTGTTTGGTGAGAAGTCTCGAGCTTCCGGGAAGATATATT 488
Db 422 CTGGAAAGTGTCAATAGCTGTTTGGTGAGAAGTCTCGAGCTTCCGGGAAGATATATT 481
QY 489 CGACTCTGTGAGAAATATTACT 510
Db 482 CGACTCTGTGAGAAATATTACT 503

RESULT 10
US-09-997-279-66
; Sequence 66, Application US/09997279
; Publication No. US20030059781A1
; GENERAL INFORMATION:
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER
; FILE REFERENCE: 210121.501C1
; CURRENT APPLICATION NUMBER: US/09/997.279
; CURRENT FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-997-279-66

Query Match 26.1%; Score 498.8; DB 10; Length 503;
Best Local Similarity 99.6%; Pred. No. 3.2e-103;
Matches 500; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 TTGCCCGTCAGACGCAACTCAGAGAATAACACAGAGAACACCAAGATTGAAACAATGGAG 68
Db 2 TCAGACCGTCAGACGCAACTCAGAGAATAACACAGAGAACACCAAGATTGAAACAATGGAG 61
QY 69 GATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCAAAAGCA 128
Db 62 GATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCAAAAGCA 121
QY 129 AGCCCAACCCAGAACCTTCTCTCTCCCATGGAGCATCTGCTCCACCATGGCCATGTC 188
Db 122 AGCCCAACCCAGAACCTTCTCTCTCCCATGGAGCATCTGCTCCACCATGGCCATGTC 181
QY 189 TACATGGGCTCCAGGGGCGAGCACCGAAGACCCAGATGGCCAAAGTGCTTCAGTTTAAATGAA 248
Db 182 TACATGGGCTCCAGGGGCGAGCACCGAAGACCCAGATGGCCAAAGTGCTTCAGTTTAAATGAA 241
QY 249 GTGGGAGCCAAATGCACTTACCCCATGACCTCCAGAGAACTTTACAGCTGTGGGTTTCATG 308
Db 242 GTGGGAGCCAAATGCACTTACCCCATGACCTCCAGAGAACTTTACAGCTGTGGGTTTCATG 301
```

```
QY 129 AGCCCAACCCAGAACCTTCTCTCTCCCATGGAGCATCTGCTCCACCATGGCCATGTC 188
Db 122 AGCCCAACCCAGAACCTTCTCTCTCCCATGGAGCATCTGCTCCACCATGGCCATGTC 181
QY 189 TACATGGGCTCCAGGGGCGAGCACCGAAGACCCAGATGGCCAAAGTGCTTCAGTTTAAATGAA 248
Db 182 TACATGGGCTCCAGGGGCGAGCACCGAAGACCCAGATGGCCAAAGTGCTTCAGTTTAAATGAA 241
QY 249 GTGGGAGCCAAATGCACTTACCCCATGACCTCCAGAGAACTTTACAGCTGTGGGTTTCATG 308
Db 242 GTGGGAGCCAAATGCACTTACCCCATGACCTCCAGAGAACTTTACAGCTGTGGGTTTCATG 301
QY 309 CAGCAGATCCAGAGGAGTAGTTATCTGATGCGATTTTGCAGGACCAAGCTGCAGATATA 368
Db 302 CAGCAGATCCAGAGGAGTAGTTATCTGATGCGATTTTGCAGGACCAAGCTGCAGATATA 361
QY 369 ATCCATTTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTATTTA 428
Db 362 ATCCATTTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTATTTA 421
QY 429 CTGGAAAGTGTCAATAGCTGTTTGGTGAGAAGTCTCGAGCTTCCGGGAAGATATATT 488
Db 422 CTGGAAAGTGTCAATAGCTGTTTGGTGAGAAGTCTCGAGCTTCCGGGAAGATATATT 481
QY 489 CGACTCTGTGAGAAATATTACT 510
Db 482 CGACTCTGTGAGAAATATTACT 503

RESULT 10
US-09-997-279-66
; Sequence 66, Application US/09997279
; Publication No. US20030059781A1
; GENERAL INFORMATION:
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER
; FILE REFERENCE: 210121.501C1
; CURRENT APPLICATION NUMBER: US/09/997.279
; CURRENT FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-997-279-66

Query Match 26.1%; Score 498.8; DB 10; Length 503;
Best Local Similarity 99.6%; Pred. No. 3.2e-103;
Matches 500; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 TTGCCCGTCAGACGCAACTCAGAGAATAACACAGAGAACACCAAGATTGAAACAATGGAG 68
Db 2 TCAGACCGTCAGACGCAACTCAGAGAATAACACAGAGAACACCAAGATTGAAACAATGGAG 61
QY 69 GATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCAAAAGCA 128
Db 62 GATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCAAAAGCA 121
QY 129 AGCCCAACCCAGAACCTTCTCTCTCCCATGGAGCATCTGCTCCACCATGGCCATGTC 188
Db 122 AGCCCAACCCAGAACCTTCTCTCTCCCATGGAGCATCTGCTCCACCATGGCCATGTC 181
QY 189 TACATGGGCTCCAGGGGCGAGCACCGAAGACCCAGATGGCCAAAGTGCTTCAGTTTAAATGAA 248
Db 182 TACATGGGCTCCAGGGGCGAGCACCGAAGACCCAGATGGCCAAAGTGCTTCAGTTTAAATGAA 241
QY 249 GTGGGAGCCAAATGCACTTACCCCATGACCTCCAGAGAACTTTACAGCTGTGGGTTTCATG 308
Db 242 GTGGGAGCCAAATGCACTTACCCCATGACCTCCAGAGAACTTTACAGCTGTGGGTTTCATG 301
```



```

; OTHER INFORMATION: EST_HUMAN HIT: BP666685.1, EVALUE 0.00e+00
US-10-029-386-13440

Query Match      20.3%; Score 387.6; DB 14; Length 596;
Best Local Similarity 99.0%; Pred. No. 6.8e-78;
Matches 390; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 902 GCTGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGACAAAAT 961
DB 203 GCAGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGACAAAAT 262

QY 962 GCTGGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGAGCATTTGA 1021
DB 263 GCTGGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGAGCATTTGA 322

QY 1022 CAGATCCATCTCTGAAAGCATGGGATGAGGAGCGCTTCAACAAGGACGGGCCAATTT 1081
DB 323 CAGATCCATCTCTGAAAGCATGGGATGAGGAGCGCTTCAACAAGGACGGGCCAATTT 382

QY 1082 CTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATG 1141
DB 383 CTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATG 442

QY 1142 GGATGTGAATGAGAGGCGCTGAAGCAGCGCTGGACAGAGGTGTTATGACAGGGAG 1201
DB 443 GGATGTGAATGAGAGGCGCTGAAGCAGCGCTGGACAGAGGTGTTATGACAGGGAG 502

QY 1202 AACTGACATGGAGGCCACAGTTGTGGCAGATCATCCGTTCTTTCTTTATTTATGCA 1261
DB 503 AACTGACATGGAGGCCACAGTTGTGGCAGATCATCCCTTTCTTTCTTTATTTATGCA 562

QY 1262 TAAGATAACCAAGTGCATTTATTTTTCGGGAGA 1295
DB 563 TAAGATAACCAAGTGCATTTATTTTTCGGGAGA 596

RESULT 13
US-10-027-632-122703/c
; Sequence 122703, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 10897.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/216,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122703
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-122703

Query Match      18.4%; Score 350.2; DB 15; Length 1256;
Best Local Similarity 99.2%; Pred. No. 3.4e-69;
Matches 352; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 902 GCTGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGACAAAAT 961
DB 355 GCAGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGACAAAAT 296

QY 962 GCTGGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGAGCATTTGA 1021
DB 295 GCTGGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGAGCATTTGA 236

QY 1022 CAGATCCATCTCTGAAAGCATGGGATGAGGAGCGCTTCAACAAGGACGGGCCAATTT 1081
DB 235 CAGATCCATCTCTGAAAGCATGGGATGAGGAGCGCTTCAACAAGGACGGGCCAATTT 176

QY 1082 CTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATG 1141
DB 175 CTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATG 116

QY 1142 GGATGTGAATGAGAGGCGCTGAAGCAGCGCTGGACAGAGGTGTTATGACAGGGAG 1201
DB 115 GGATGTGAATGAGAGGCGCTGAAGCAGCGCTGGACAGAGGTGTTATGACAGGGAG 56

QY 1202 AACTGACATGGAGGCCACAGTTGTGGCAGATCATCCGTTCTTTCTTTATTTAT 1256
DB 55 AACTGACATGGAGGCCACAGTTGTGGCAGATCATCCCTTTCTTTCTTTATTTAT 1

RESULT 14
US-10-029-386-27140
; Sequence 27140, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; POLYMORPHISM ANALYSIS TWO
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27140
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR18.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.32
; OTHER INFORMATION: EST_HUMAN HIT: BP666685.1, EVALUE 0.00e+00
; OTHER INFORMATION: NT HIT: M24657.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P05120, EVALUE 3.00e-47
US-10-029-386-27140

Query Match      18.0%; Score 343.2; DB 14; Length 354;
Best Local Similarity 99.1%; Pred. No. 6.3e-68;
Matches 345; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 902 GCTGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGACAAAAT 961
DB 7 GCAGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGACAAAAT 66

QY 962 GCTGGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGAGCATTTGA 1021
DB 67 GCTGGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGAGCATTTGA 126

QY 1022 CAGATCCATCTCTGAAAGCATGGGATGAGGAGCGCTTCAACAAGGACGGGCCAATTT 1081
DB 127 CAGATCCATCTCTGAAAGCATGGGATGAGGAGCGCTTCAACAAGGACGGGCCAATTT 186

QY 1082 CTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATG 1141
DB 187 CTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATG 246

QY 1142 GGATGTGAATGAGAGGCGCTGAAGCAGCGCTGGACAGAGGTGTTATGACAGGGAG 1201

```

247 GGATGTGAATGAGGAGGACCTGAGCAGCCGCTGCAACAGGAGGTGTATGACAGGGAG 306
1202 AACTGGACATGAGGCCACACAGTTTGTGGCAGATCATCCGTTTCTTTT 1249
307 AACTGGACATGAGGCCACACAGTTTGTGGCAGATCATCTTTTCTTTT 354

RESULT 15
JS-10-029-386-865/c
; Sequence 865, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 865
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009802.13
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.84
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.33
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: NT HIT: M1551.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF66685.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P05120, EVALUE 3.00e-45
US-10-029-386-865

Query Match 17.7%; Score 337.2; DB 14; Length 544;
Best Local Similarity 99.1%; Pred. No. 1.9e-66;
Matches 339; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 902 GCTGCTGGAAGTGAATTAACCTATGACAACTCAACAGTGGACCAAGCAAGCAAAAT 961
Db 342 GCAGCTGGAAGTGAATTAACCTATGACAACTCAACAGTGGACCAAGCAAAAT 283
Qy 962 GCCTGAAGATGAAGTTAGGTATACATACCCAGTTCAAATTAGAAGACATTATGAAT 1021
Db 282 GCCTGAAGATGAAGTTAGGTATACATACCCAGTTCAAATTAGAAGACATTATGAAT 223
Qy 1022 CAGATCCATTCTGAAAGCATGGGATGGAGGACGCTTCAACAAAGGACGGGCCAATTT 1081
Db 222 CAGATCCATTCTGAGNAGCATGGGATGGAGGACGCTTCAACAAAGGACGGGCCAATTT 163
Qy 1082 CTCAGGGATGTCGGAGAGGAATGACCTGTTCTTCTGAAGTGTCCACCAAGCCATGGT 1141
Db 162 CTCAGGGATGTCGGAGAGGAATGACCTGTTCTTCTGAAGTGTCCACCAAGCCATGGT 103
Qy 1142 GGATGTGAATGAGGAGGGCACTGAAGCAGCGCGTGGCAGAGGAGGTGTATGACAGGGAG 1201
Db 102 GGATGTGAATGAGGAGGGCACTGAAGCAGCGCGTGGCAGAGGAGGTGTATGACAGGGAG 43
Qy 1202 AACTGGACATGAGGCCACACAGTTTGTGGCAGATCATCCGTT 1243
Db 42 AACTGGACATGAGGCCACACAGTTTGTGGCAGATCATCTTTT 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DN nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:18:38 ; Search time 3158.9 Seconds
(without alignments)
18036.975 Million cell updates/sec

Title: US-10-084-817-27

Perfect score: 1908
Sequence: 1 gagagattccgcgcaga.....aacaaaaataaaaaaag 1908

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:

1: em_estba.*

2: em_estba.*

3: em_estin.*

4: em_estin.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hcc.*

9: gb_est1.*

10: gb_est2.*

11: gb_hcc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rtd.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gsl.*

29: gb_gsl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1241.6	65.1	1248	29	AY416313	AY416313 Homo sapi
2	1046	54.8	1080	29	AY416314	AY416314 Pan trogl
3	1041.4	54.6	1145	9	AL550163	AL550163 AL550163
4	1027.8	53.9	1126	9	AL574114	AL574114 AL574114

5	1006.4	52.7	1965	11	AK081487	AK081487 Mus muscu
6	1000.8	52.5	1952	11	AK090049	AK090049 Mus muscu
7	970.4	50.9	1201	9	AL579236	AL579236 AL579236
8	947.6	48.7	1044	9	AL552817	AL552817 AL552817
9	932.6	48.9	1044	9	AL552852	AL552852 AL552852
10	931	48.8	991	9	AL576638	AL576638 AL576638
11	925.4	48.5	1000	9	AL553610	AL553610 AL553610
12	913	47.9	1041	13	BX396200	BX396200 BX396200
13	888	46.5	1201	9	AL557133	AL557133 AL557133
14	884.6	46.4	959	9	AL544520	AL544520 AL544520
15	883.4	46.3	989	9	AL545835	AL545835 AL545835
16	880	46.1	1040	13	BX338884	BX338884 BX338884
17	873.6	45.8	1017	9	AL545435	AL545435 AL545435
18	873.4	44.9	898	9	AL547037	AL547037 AL547037
19	856.4	44.8	1248	29	AY416315	AY416315 Mus muscu
20	854.4	44.8	1248	29	AY416315	AY416315 Mus muscu
21	840.8	44.1	1006	9	AL572149	AL572149 AL572149
22	839.8	44.0	1201	13	BX424501	BX424501 BX424501
23	834.8	43.8	901	9	AL544761	AL544761 AL544761
24	831	43.6	903	9	AL545468	AL545468 AL545468
25	805.6	42.2	833	9	AL573314	AL573314 AL573314
26	797	41.8	1201	9	AL570830	AL570830 AL570830
27	769.6	40.3	865	9	AU138758	AU138758 AU138758
28	769.6	40.3	900	13	BX424585	BX424585 BX424585
29	758.4	39.7	771	14	CD742955	CD742955 UI-H-F11
30	755.4	39.6	764	14	CD521331	CD521331 AGENCOURT
31	753.8	39.5	855	9	AU139494	AU139494 AU139494
32	753.6	39.5	764	14	CA313606	CA313606 UI-CF-FNO
33	753	39.5	818	9	AL557586	AL557586 AL557586
34	750.8	38.4	792	9	AU139617	AU139617 AU139617
35	748.2	38.3	1201	9	AL544856	AL544856 AL544856
36	748.2	39.2	754	13	BX099033	BX099033 BX099033
37	747.4	38.8	779	14	CA310353	CA310353 UI-H-F11
38	740.8	38.7	756	14	CA748486	CA748486 UI-H-F11
39	739.2	38.7	894	14	CD522880	CD522880 AGENCOURT
40	738	38.7	980	9	AL571400	AL571400 AL571400
41	738	38.7	980	9	AL571400	AL571400 AL571400
42	733.4	38.4	1201	13	BX381897	BX381897 BX381897
43	733.2	38.4	760	14	CD364488	CD364488 UI-H-F11
44	733.2	38.4	760	14	CD364488	CD364488 UI-H-F11
45	732.8	38.4	755	14	CD743065	CD743065 UI-H-F11

ALIGNMENTS

RESULT 1	AY416313	1248 bp	DNA	linear	GSS 17-DEC-2003
LOCUS	AY416313	Homo sapiens SERPINB2 gene, VIRTUAL TRANSCRIPT, partial sequence,			
DEFINITION	AY416313.1	GI:39772273			
ACCESSION	AY416313				
VERSION	AY416313.1				
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1248)				
AUTHORS	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.O., Adams, M.D. and Cargill, M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 1248)				
AUTHORS	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.O., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..1248
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..>1248
/gene="SERPINB2"
/locus_tag="HCM5849"
ORIGIN
Query Match 65.1%; Score 1241.6; DB 29; Length 1248;
Best Local Similarity 99.7%; Pred. No. 1.7e-221;
Matches 1244; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 63 ATGGAGGATCTTTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGGCA 122
DB 1 ATGGAGGATCTTTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGGCA 60
QY 123 AAAGCAAGCCCAACCAAGACCTCTCTCCCATGGAGCATCTGTCACCAATGGCC 182
DB 51 AAAGCAAGCCCAACCAAGACCTCTCTCCCATGGAGCATCTGTCACCAATGGCC 120
QY 183 ATGGTCTACATGGGCTCCAGGGGAGCAGCAGCAAGACAGATGGCAGAGGTGCTTCAGTTT 242
DB 121 ATGGTCTACATGGGCTCCAGGGGAGCAGCAGCAAGACAGATGGCAGAGGTGCTTCAGTTT 180
QY 243 AATGAAGTGGAGCCCAATGAGTATCCCATGATGCTCAGAGCAATTTACAGCTGTGG 302
DB 181 AATGAAGTGGAGCCCAATGAGTATCCCATGATGCTCAGAGCAATTTACAGCTGTGG 240
QY 303 TTATGATGAGCATCCAGAGGGGTAGTTATCTCTGATGCGATTTTGCAGGCAACAGCTGCA 362
DB 241 TTATGATGAGCATCCAGAGGGGTAGTTATCTCTGATGCGATTTTGCAGGCAACAGCTGCA 300
QY 363 GATGAATGCAATTCATCTTCGGCTCTCAGCTCTGCAATCAATGATCCACAGGGAAT 422
DB 301 GATGAATGCAATTCATCTTCGGCTCTCAGCTCTGCAATCAATGATCCACAGGGAAT 360
QY 423 TATTTACTGGAAAGTGTCAATAAGCTGTTTGGTGAGAAGTCTGCGAGCTTCGGGAAGAA 482
DB 361 TATTTACTGGAAAGTGTCAATAAGCTGTTTGGTGAGAAGTCTGCGAGCTTCGGGAAGAA 420
QY 483 TATTTACTGGTCTGTCAGAAATATTACTCTCAGACCCAGGAGTGTCTCTAGAA 542
DB 421 TATTTACTGGTCTGTCAGAAATATTACTCTCAGACCCAGGAGTGTCTCTAGAA 480
QY 543 TGTGCAAGAAAGCTAGAAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCAAA 602
DB 481 TGTGCAAGAAAGCTAGAAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCAAA 540
QY 603 ATCCCAAACTGTGTACCTGAAGGTTCTGTAGATGGGATACAGGATGGTCTCTGGTGAAT 662
DB 541 ATCCCAAACTGTGTACCTGAAGGTTCTGTAGATGGGATACAGGATGGTCTCTGGTGAAT 600
QY 663 GTGTCTACTTCAAGAAAGTGTGAAGAACTCAATTTGAGAGAACTAAATGGGCTTTAT 722
DB 601 GTGTCTACTTCAAGAAAGTGTGAAGAACTCAATTTGAGAGAACTAAATGGGCTTTAT 660
QY 723 CTTTCCGTGTAACTCGGCTCAGCGACACCTGTACAGATGATGTACTTGGTCAAAAG 782
DB 661 CTTTCCGTGTAACTCGGCTCAGCGACACCTGTACAGATGATGTACTTGGTCAAAAG 720
QY 783 CTAAACATTTGGATACATAGAGACCTAAGGCTCAGATTTAGAACTCCCATATGCTGGA 842
DB 721 CTAAACATTTGGATACATAGAGACCTAAGGCTCAGATTTAGAACTCCCATATGCTGGA 780
QY 843 GATGTTAGCATGTTCTTGTGCTTCCAGATGAAATTCGGATGTGTCACCTGGCTTGGAG 902
DB 781 GATGTTAGCATGTTCTTGTGCTTCCAGATGAAATTCGGATGTGTCACCTGGCTTGGAG 840

QY 903 CTGCTGGAAGTGCAATTAACCTATGACAACTCAACAGTGGAGCCAGCAAGACAAAATG 962
DB 841 CTGCTGGAAGTGCAATTAACCTATGACAACTCAACAGTGGAGCCAGCAAGACAAAATG 900
QY 963 GCTGAAGATGAAGTGTAGGTATATACATACCCAGTTCAAATTTAGAGAGCATTTGAACATC 1022
DB 901 GCTGAAGATGAAGTGTAGGTATATACATACCCAGTTCAAATTTAGAGAGCATTTGAACATC 960
QY 1023 AGATCCATTTCTGAAAAGCATGGGCATGGAGAGCCCTTCAACAGGGGACGGGCCAATTC 1082
DB 961 AGATCCATTTCTGAGAAGCATGGGCATGGAGAGCCCTTCAACAGGGGACGGGCCAATTC 1020
QY 1083 TCAGGATGTCGAGAGAGGAATGACCTGTTCTTCTGAAGTGTTCACCAAGCCCATGGT 1142
DB 1021 TCAGGATGTCGAGAGAGGAATGACCTGTTCTTCTGAAGTGTTCACCAAGCCCATGGT 1080
QY 1143 GATGTCAATGAGAGGGCACTGAAGCAGCGCTGGCAGAGGTGTTATGACAGGGAGA 1202
DB 1081 GATGTCAATGAGAGGGCACTGAAGCAGCGCTGGCAGAGGTGTTATGACAGGGAGA 1140
QY 1203 ACTGACATGGAGGCCACAGTTTGTGGCAGATCATCGTTTCTTTCTTATTTATGCA 1262
DB 1141 ACTGACATGGAGGCCACAGTTTGTGGCAGATCATCGTTTCTTTCTTATTTATGCA 1200
QY 1263 AAGATAACCAAGTGCATTTTATTTTTCGGCAGATTTTCTCACCCTAA 1310
DB 1201 AAGATAACCAAGTGCATTTTATTTTTCGGCAGATTTTCTCACCCTAA 1248
RESULT 2
AY416314 1080 bp DNA linear GSS 17-DSC-2003
LOCUS Pan troglodytes SERPINB2 gene, VIRUAL TRANSCRIPT, partial
DEFINITION sequence, genomic survey sequence.
ACCESSION AY416314
VERSION AV416314.1 GI:39772274
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 1080)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1080)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..1080
Location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene <1..>1080
/gene="SERPINB2"
/locus_tag="HCM5849"
ORIGIN
Query Match 54.8%; Score 1046; DB 29; Length 1080;
Best Local Similarity 97.7%; Pred. No. 5.2e-185;

	Matches	1055;	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;
231	GTGCTT	CAGTTA	TGAGTGGAGCC	AAATGAGTTAC	CCCCATGACTCC	CAGAGAACTTT	290			
1	GTGCTG	CAGTTA	TGAGTGGAGCC	AAATGAGTTAC	CCCCATGACTCC	CAGAGAACTTT	60			
291	ACCAAGCT	GTGGTT	CATGAGAGAT	CCAGAGGAGT	TATCTCTGATGGATTT	GCAG	350			
61	ACCAAGCT	GTGGTT	CATGAGAGAT	CCAGAGGAGT	TATCTCTGATGGATTT	GCAG	120			
351	GCACAGCT	GCAGAT	TAATAATCA	ATTCCTTCCGCTCT	CTCTAGCTCTGCAATCA	ATGCA	410			
121	GCACAGCT	GCAGAT	TAATAATCA	ATTCCTTCCGCTCT	CTCTAGCTCTGCAATCA	ATGCA	180			
411	TCCACAGG	GAATTA	TACTGGAAGT	GTCAATAGCTGTT	GGTGAGAGTCTGGGAGC	470				
181	TCCACAGG	GAATTA	TACTGGAAGT	GTCAATAGCTGTT	GGTGAGAGTCTGGGAGC	240				
471	TTCGGG	AGAAAT	TATTCGACT	CTGTGCAAAATAT	TACTCTCAGAACCC	CCAGGAGTA	530			
241	TTCGGG	AGAAAT	TATTCGACT	CTGTGCAAAATAT	TACTCTCAGAACCC	CCAGGAGTA	300			
531	GACTTCT	CTAGAT	GTGCAAGCT	GAGAGGATTA	ATTCCTGGGTCAAGACT	CAA	590			
301	GACTTCT	CTAGAT	GTGCAAGCT	GAGAGGATTA	ATTCCTGGGTCAAGACT	CAA	360			
591	ACCAAGC	CAAAAT	CCAACTT	TACTGAGTTC	TCTGTAGATGGGAT	CAGGATG	650			
361	ACCAAGC	CAAAAT	CCAACTT	TACTGAGTTC	TCTGTAGATGGGAT	CAGGATG	420			
651	GTCCTGT	GAATGCT	GTCTACT	TCAAAGGAAAGT	GAAAACTCCATTTG	AGAGAACTA	710			
421	GTCCTGT	GAATGCT	GTCTACT	TCAAAGGAAAGT	GAAAACTCCATTTG	AGAGAACTA	480			
711	AATGGCT	TATATCT	TTCGTTAACT	CGGTCAAGT	CGGACACCTGT	CAGATGATGAC	770			
481	AATGGCT	TATATCT	TTCGTTAACT	CGGTCAAGT	CGGACACCTGT	CAGATGATGAC	540			
771	TTGGGT	GAAAAGCT	AAACAT	TGATAGAGAC	CTAAAGGCTCAGAT	TCTAGAACTC	830			
541	TTGGGT	GAAAAGCT	AAACAT	TGATAGAGAC	CTAAAGGCTCAGAT	TCTAGAACTC	600			
831	CCATATG	CTGAGAT	GTGTAGCAT	GTCTTGTGCTT	CCAGATGAAT	TGCGATGTGTC	890			
601	CCATATG	CTGAGAT	GTGTAGCAT	GTCTTGTGCTT	CCAGATGAAT	TGCGATGTGTC	660			
891	ACTGGCT	TGAGCT	GTCTGAAAAGT	GAATAACCTAT	GACAACTCAAC	AGTGGACGAGC	950			
661	ACTGGCT	TGAGCT	GTCTGAAAAGT	GAATAACCTAT	GACAACTCAAC	AGTGGACGAGC	720			
951	AAAGCA	AAATGGCT	GAAGATGA	AGTTGAGGTAT	ATATACCCCACTCA	ATTAAGAGAG	1010			
721	AAAGCA	AAATGGCT	GAAGATGA	AGTTGAGGTAT	ATATACCCCACTCA	ATTAAGAGAG	780			
1011	CATTATGA	CACTCAGAT	CCATCTG	AAAAGCATGGG	CATGGAGACGCTT	CAACAGGGA	1070			
781	CATTATGA	CACTCAGAT	CCATCTG	AAAAGCATGGG	CATGGAGACGCTT	CAACAGGGA	840			
1071	CGGGCA	ATTTCTCAGG	ATGTGCGAG	GAATGACCTGTT	CTCTGAGTGTCCAC		1130			
841	CGGGCA	ATTTCTCAGG	ATGTGCGAG	GAATGACCTGTT	CTCTGAGTGTCCAC		900			
1131	CAAGCAT	GTGGATGT	GAATGAGG	AGGCACTGAAG	CAAGCCGCTGGCA	CAGAGGTGTT	1190			
901	CAAGCAT	GTGGATGT	GAATGAGG	AGGCACTGAAG	CAAGCCGCTGGCA	CAGAGGTGTT	960			
1191	ATGACAGG	AGAACTGG	ACATGGAG	CCACAGTTT	GTGCGAGATCAT	CTGTTCTTTT	1250			
961	ATGACAGG	AGAACTGG	ACATGGAG	CCACAGTTT	GTGCGAGATCAT	CTGTTCTTTT	1020			
1251	CTTATTAT	GCATAA	CAAGTGCAT	TTTATTTT	GGCGAGATTT	TGCTCACCTAA	1310			
1021	CTTATTAT	GCATAA	CAAGTGCAT	TTTATTTT	GGCGAGATTT	TGCTCACCTAA	1080			

RESULT 3
AL550163
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL550163 1145 bp mRNA linear EST 31-MAY-2003
AL550163 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS01040YG06 5-PRIME, mRNA sequence.
AL550163
AL550163.2 GI:31271981
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1145)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12886866.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9202.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgt-bn/cluster.cgi?seq=CS01040BD03QPI&cluster=9202.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS01040BD03QPI.
Location/Qualifiers
1..1145
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01040YG06"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source

1..1145
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01040YG06"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 54.6%; Score 1041.4; DB 9; Length 1145;
Best Local Similarity 98.2%; Pred. No. 3.7e-184;
Matches 1066; Conservative 7; Mismatches 9; Indels 3; Gaps 2;

QY	10	TGCCCGT	CAGACAG	CAACTCT	CCCTCC	CCATGG	AGCATCT	CGTCC	ACCATG	GCCTG	GGTCT	189
Db	64	TGCCCGT	CAGACAG	CAACTCT	CCCTCC	CCATGG	AGCATCT	CGTCC	ACCATG	GCCTG	GGTCT	243
QY	70	ATCTTTG	TGTGG	CAACAC	ACTCT	TTTGC	CTCAAT	TTATT	CAAGCAT	CTG	CAAAAGCA	129
Db	124	ATCTTTG	TGTGG	CAACAC	ACTCT	TTTGC	CTCAAT	TTATT	CAAGCAT	CTG	CAAAAGCA	183
QY	130	GCCTCC	ACCCAG	AACTCT	TTCTCT	CCCTCC	ATGG	AGCATCT	CGTCC	ACCATG	GCCTG	189
Db	184	GCCTCC	ACCCAG	AACTCT	TTCTCT	CCCTCC	ATGG	AGCATCT	CGTCC	ACCATG	GCCTG	243
QY	190	ACATGG	CTCCAG	GGGAG	CAGC	CGGAA	GACAG	ATG	CGC	CAAGG	TCTT	249
Db	244	ACATGG	CTCCAG	GGGAG	CAGC	CGGAA	GACAG	ATG	CGC	CAAGG	TCTT	303
QY	250	TGGAGC	CAATC	AGTTAC	CCCC	CCATG	CTCC	AGAACT	TTTAC	CAGCT	TGTGG	309
Db	304	TGGAGC	CAATC	AGTTAC	CCCC	CCATG	CTCC	AGAACT	TTTAC	CAGCT	TGTGG	363
QY	310	AGCAG	ATCC	AGAGG	GTAGT	TATCT	CTG	ATG	CG	CGAT	TTT	369
Db	364	AGCAG	ATCC	AGAGG	GTAGT	TATCT	CTG	ATG	CG	CGAT	TTT	423

370 TCATTCATCCCTTCGCTCTCAGCTCTGCAATCAATGATCCACAGGGAATATTATTC 429
424 TCATTCATCCCTTCGCTCTCAGCTCTGCAATCAATGATCCACAGGGAATATTATTC 483
430 TGGAAAGTGTCAATAAGCTCTTGGTGAGAGTCTGCGAGCTTCCGGGAGGAATATTC 489
484 TGGAAAGTGTCAATAAGCTCTTGGTGAGAGTCTGCGAGCTTCCGGGAGGAATATTC 543
490 GACTCTGTGAGAAATATTACTCTCTCAGAACCCAGGAGTACTCTTCTAGATGTGAG 549
544 GACTCTGTGAGAAATATTACTCTCTCAGAACCCAGGAGTACTCTTCTAGATGTGAG 603
550 AAGAAGCTAGAAAAGAAATTAATTCCTGGTCAAGACTCAAAACCAAGGCAAAATCCCAA 609
604 AAGAAGCTAGAAAAGAAATTAATTCCTGGTCAAGACTCAAAACCAAGGCAAAATCCCAA 663
610 ACTTGTTACTCTGAAGTCTCTGATAGGAGATCCAGGATGGTCTGCTGGTGAATGCTGCT 669
664 ACTTGTTACTCTGAAGTCTCTGATAGGAGATCCAGGATGGTCTGCTGGTGAATGCTGCT 723
670 ACTTCAAGGAAAGTGGAAAATCTCAATTGAGAGAAACTAAATGGGCTTTATCCTTTCC 729
724 ACTTCAAGGAAAGTGGAAAATCTCAATTGAGAGAAACTAAATGGGCTTTATCCTTTCC 783
730 GTCTAAACTCGGCTCAGCGCACCTCTGACAGATGATGCTGCTGAAAGCTAAACA 789
784 GTGTAATCTCGGCTCAGCGCACCTCTGACAGATGATGCTGCTGAAAGCTAAACA 843
790 TTGGATACATAGAGAGCTTAAAGGCTCAGATTCCTAGAACTCCATATGCTGGAGATGTA 849
844 TTGGATACATAGAGAGCTTAAAGGCTCAGATTCCTAGAACTCCATATGCTGGAGATGTA 903
850 GCATGTTCTTGTGCTTCCAGATGAATGCGGATGCTGCTGAGGCTGCTGG 909
904 GCATGTTCTTGTGCTTCCAGATGAATGCGGATGCTGCTGAGGCTGCTGG 963
910 AAGTGAATTAACCTATGACAACTCAAAAGTGGAGCCAGAGAAAGAAATGGCTGAAG 969
964 AAGTGAATTAACCTATGACAACTCAAAAGTGGAGCCAGAGAAAGAAATGGCTGAAG 1023
970 ATGAAGTTGAGGTATACATACCCAGTCAAAATGAGAGAGATTAAGAACTCAGATCCA 1029
1024 ATGAAGTTGAGGTATACATACCCAGTCAAAATGAGAGAGATTAAGAACTCAGAT-CA 1082
1030 TTCTGAAAAGCATGGGATGAGGAGCGCTTCAACAGGAGCGGCCAAATTTCTAGGGA 1089
1083 TTCTGAAAAGCATGGGATGAGGAGCGCTTCAACAGGAGCGGCCAAATTTCTAGGGA 1140
1090 TGTCG 1094
1141 GKYG 1145

RESULT 4
AL574114/c 1126 bp mRNA linear EST 31-MAY-2003
LOCUS AL574114 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1040YG06 3-PRIME, mRNA sequence.
ACCESSION AL574114
VERSION 1
KEYWORDS EST.
SOURCE AL574114.2 GI:31295449
ORGANISM Homo sapiens (human)
REFERENCE Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1126)
Li, W.B., Gruber, C., Jessup, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:12934006.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9202.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1040BD03NP1&cluster=9202.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0D1040BD03NP1.
FEATURES
Location/Qualifiers
1..1126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1040YG06"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 53.9%; Score 1027.8; DB 9; Length 1126;
Best Local Similarity 98.1%; Pred. No. 1.3e-181;
Matches 1056; Conservative 13; Mismatches 5; Indels 3; Gaps 3;
753 CCGTACAGATGATGATCTCCGCGAAAGCTAAACATTTGGATACATAGAACCTAAAG 812
1111 MCTGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1054
813 GCTCAGATTTAGAACTCCCATATGCTGGAGATGATGATGATGATGATGATGATGATGAT 872
1053 GCTCAGATTTAGAACTCCCATATGCTGGAGATGATGATGATGATGATGATGATGATGAT 994
873 GAATTTGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 932
993 GAATTTGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 934
933 CTCAACAAGTGGAGCCAGCAAAAGCAAAATGCTGAAGATGAAGTTGAGGTATACATACCC 992
933 CTCAACAAGT-GACCAGCAAAAGCAAAATGCTGAAGATGAAGTTGAGGTATACATACCC 875
993 CAGTTCAATTAGAGAGCATTTAGAACTCAGATCCATTTCTGAAAGCATGGCATGGAG 1052
874 CAGTTCAATTAGAGAGCATTTAGAACTCAGATCCATTTCTGAAAGCATGGCATGGAG 815
1053 GACGCTTCAACAAGGAGCGGCCAAATTTCTCAGGATGTCGAGAGGAATGACCTGTTT 1112
814 GACGCTTCAACAAGGAGCGGCCAAATTTCTCAGGATGTCGAGAGGAATGACCTGTTT 755
1113 CTTTCTGAAGTGTTCACCAAGCCATGCTGGATGTAATGAGGAGGCGCTGAAGCAGCC 1172
754 CTTTCTGAAGTGTTCACCAAGCCATGCTGGATGTAATGAGGAGGCGCTGAAGCAGCC 695
1173 GCTGGCAAGAGGTGTTATGACAGGAGAACTGGACATGGAGGCCACAGTTTGTGGCA 1332
694 GCTGGCAAGAGGTGTTATGACAGGAGAACTGGACATGGAGGCCACAGTTTGTGGCA 635
1233 GATCATCGCTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1292
634 GATCATCGCTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 575
1293 AGATTTGCTCACCCCTCAAACTAAGCGGTGCTGCTTCTGCAAAAGATTTTGTAGATGAGC 1352
574 AGATTTTCTCACCCCTCAAACTAAGCGGTGCTGCTTCTGCAAAAGATTTTGTAGATGAGC 515
1353 TGTGTGCTCAGAAATGCTATTTCAAAATGCAAAATTTAGAGATGTTTCTTACATATT 1412
514 TGTGTGCTCAGAAATGCTATTTCAAAATGCAAAATTTAGAGATGTTTCTTACATATT 455
1413 TGTGTGCTCAGAAATGCTTCTGCTACCCATTAATTAATAACACAGAAATATTAGACAA 1472

db	454	TCCTCTCTTCTGAACWACVTCYCTACCCACTAAATAAAAACACAGAAATAATTAGACAA	395
2y	1473	TTGCTCTATTATAACATGACCAACCTTATTATCATTTGGTCTTCTTAAATGGATCATGCC	1532
db	394	TTGCTCTATTATAACATGACCAACCTTATTATCATTTGGTCTTAAATGGATCATGCC	335
2y	1533	CATTAGATTTTCCCTACTATCATCAGTTTATTTTATAACATTAACCTTTTACTTTGTTATTT	1592
db	334	CATTAGATTTTCCCTACTATCATCAGTTTATTTTATAACATTAACCTTTTACTTTGTTATTT	275
2y	1593	ATTATTTTATAAATGGTGAGTTTTTAAATATTATGCTACTGCTCTTAAATGGATCATGCC	1652
db	274	ATTATTTTATAAATGGTGAGTTTTTAAATATTATGCTACTGCTCTTAAATGGATCATGCC	215
2y	1653	TAAAGTTTATAGAAGCAGATGATCTGTTAATTTTCTATCTAATAAATGCTTTAATTTGTT	1712
db	214	TAAAGTTTATAGAAGCAGATGATCTGTTAATTTTCTATCTAATAAATGCTTTAATTTGTT	155
2y	1713	TCATAATGAAGTAAGTAGGATATCCCTCATGCTCTTCTGTAATAAATCTGGAATAA	1772
db	154	TCATAATGAAGTAAGTAGGATATCCCTCATGCTCTTCTGTAATAAATCTGGAATAA	95
2y	1773	ACATTAAACATAGGCAATATATGTTATGCTGATTTCTAGAAATACATACACATA	1829
db	94	ACATTAAACATAGGCAATATATGTTATGCTGATTTCTAGAAATACATACACATA	38
RESULT 5			
LOCUS	AK081487		
DEFINITION	Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:Cl30020L09 product:serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 2, full insert sequence.		
ACCESSION	AK081487		
VERSION	AK081487.1	GI:26349182	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Mech. Enzymol.	303, 19-44 (1999)	
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitesunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakeguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Chara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res.	10 (11), 1757-1771 (2000)	
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitesunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakeguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Chara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature	409, 685-690 (2001)	
REFERENCE	5		
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature	420, 563-573 (2002)	
REFERENCE	6 (bases 1 to 1965)		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)		
COMMENT	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.		
FEATURES	Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/		
source	Location/Qualifiers		
	1..1965		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="FANTOM DB:Cl30020L09"		
	/db_xref="MGI:2413892"		
	/db_xref="taxon:10090"		
	/clone="Cl30020L09"		
	/tissue_type="head"		
	/clone_lib="RIKEN full-length enriched mouse cDNA library"		
	/dev_stage="16 days embryo"		
	67..1314		
	/note="unnamed protein product; putative serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 2 (MGD MGI:97609, GB NM_011111, evidence: BLASTN, 100%, match=1899)"		
	/codon_start=1		
	/protein_id="BAC38231.1"		
	/db_xref="GI:26349183"		
	/translation="MEELSMANTFALNLKQIEKSNSTONIFSPWISSTLAIVLL IHSATFQQMAKVLQFNEIGSYGITRNPENFGCDPAQIQENYPAIQAQADK IHSATFSSSTINTPGDYLLESANKLFGKSAKFKEEYIQLSKYFTEPAFLE CAEAREKINSWVKTOTKEIPNLLPEGSVDKTMVLVNAVYFKWKTPEKKING LYPFVNSHESIPVOMFLHAKNIGVIKDLKQILBLPTGNTSMILLLLPDEIDAS TGLELLESINFANFNKWSKDTLDDVVVLPKFKLAQSVELKSIQSMGMEIDAFN KGNFNSGMSERNDLFLFVGRSPSP"		
	PFLFFIMDKTHTLFLVGRSPSP		
	1948..1953		
	/note="putative"		
	1965		
	/note="putative"		
	polyA_signal		
	polyA_site		
	ORIGIN		
	Query Match 52.7%; Score 1006.4; DB 11; Length 1965;		
	Best Local Similarity 73.5%; Pred. No. 1.1e-177;		
	Matches 1362; Conservative 0; Mismatches 461; Indels 29; Gaps 5;		
	QY 40 CAGAGAACACACAGATTGAACAAATGGAGGATCTTTGTGTGGCAACACACTCTTTGCC 99		

Db 44 CAGAGATACGAGATGAAACAAATGGAAGACTTTCCATGGCAACACCAATGTTTGCCC 103
Qy 100 TCAATTTATTCAAGCATCTGGCAAAAGCAAGCCACCCAGAACTCTCTCTCCGCCAT 159
Db 104 TCAATCTCTTAAGCAGATAGAAAAATCAAACCTCTAGCCAGAACATCTTTATCTCTCCAT 163
Qy 160 GGAGCATCTCGTCCACCATGGCCATGCTACATGGGCTCCAGGGGAGCAGCCGAAGACC 219
Db 164 GGAGCATCTCATCAACATTTGGCCATAGTTCTCTCGGTGCTGGGGGTAACACTGNACAGC 223
Qy 220 AGATGCCCAAGGTCCTTCAGTTTAATGAAGTGGAGGCCAATGCAAGTACCCGCCATGACTC 279
Db 224 AGATGCCCAAGGTCCTTCAGTTTAATGAATGGCAGTTATGGTATCAACCAAGAAACC 283
Qy 280 CAGAGAACTTTACAGCTGTGGGTTTCATGAGCAGATCCAGAAAGGTAGTTATCTCGATG 339
Db 284 CAGAGAACTTCAGTGGCTGTGATTTCCGACACAGATACAGAAAGAAATTTATCTAGTG 343
Qy 340 CGATTTTGAGGCACCAAGCTGCGAGATAAATGCCATTCATCTCCGGCTCTCTAGCTCTG 399
Db 344 CTATTTTACAGGCACCAAGCAGAGATAAATGCCATTCAGCCCTCTCTCTCTAGCTCAA 403
Qy 400 CAATCAATGCATCCACAGGGAATTTATTTACTGGAAGTGTCAATAAGCTGTTTGGTGAGA 459
Db 404 CAATCAACACACACAGGGGATTTATTTGTAGAAAGTGCAACAGCTGTTGGAGAGA 463
Qy 460 AGTCTCGAGCTTCGGGAGAGATATATTCGACTCTGTGAGAAATATTTACTCTCAGAAC 519
Db 464 AGTCTCGAAGATTCAGAGAGAGATATATTCGAACTCTCTAAGAAATATTTACTCAACAGAAC 523
Qy 520 CCAGGAGTGTCTCTAGATGTCAGAGAGAGCTAGAGAAAGATTAATCTCTGGG 579
Db 524 CAGAGCAGTGGACTTCCTTGAATGTGCTGAGAGCTAGGGAAGATTAATCTCTGGG 583
Qy 580 TCAAGACTCAACCAAGGCAAAATCCCAACTTGTACTCTGAGAGGTTCTGTAGATGGG 639
Db 584 TCAAGACTCAACCAAGGCAAAATCCCAACTGTCTACCCGAAAGTTCTGTAGATGAAG 643
Qy 640 ATACCAAGGATGTCCTGTGAAAGTGTCTACTTCAAGAGAAAGTGGAAACTCCATTTG 699
Db 644 ACACCAAGATGGTGTGCTGTGAAAGTGTCTACTTCAAGAGAAAGTGGAAACTCCATTTG 703
Qy 700 AGAAGAACTPAATAGGCTTTATCTCTTCCGTGAAAGTGTCTGAGAGGCTGAGCTGAC 759
Db 704 AGAAGAACTTAATAGGCTTTATCTCTTCCGTGAAAGTGTCTGAGAGGCTGAGCTGAC 763
Qy 760 AGATGATGTACTTGGTGAAGAGCTAAACATTTGGATACATAGAGACCTAAGGCTCAGA 819
Db 764 AGATGATGTCTTCCATGCAAGAGCTGAACATTTGGATACATTAAGGACCTGAGACTCAGA 823
Qy 820 TTCTAGAACTCCCATATGCTGGAGATGTTAGCATGTTCTTGTGCTTCCAGATGAATTTG 879
Db 824 TCCTAGAACTTCGGCATACTGGAACATCTAGCATGCTCTGTTGCTTCCGATGAGATTG 883
Qy 880 CGATGTGCTCACTGGCTTGGAGCTGTGGAAGTGAATTAACCTATGACAAACTCAACA 939
Db 884 AGACCGATCCATCTGCTTGAATTTGCTGGAAGTGAATTAACCTTGGCAACTTCAACA 943
Qy 940 AGTGGACCAAGCAAGCAAAATGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCA 999
Db 944 AGTGGATCAGCAAGACACACTGGATGAAGATGATGTTTGTGCTTACATTTCCCAAGTTCA 1003
Qy 1000 AATTAGAAGAGCATTAATGAATCAGATCCATCTGAAAAGATGGCATGGAGGACGCT 1059
Db 1004 AACTGGCACAAGCTACGAACTCAAGTCCATTTCTCAAGCATGGCATGGAGATGCT 1063
Qy 1060 TCAACAGGACGGGCCAAATTTCTCAGGGATGTGAGAGGAATGACCTGTTCTTCTG 1119
Db 1064 TCAACAGGACGGGCCAAATTTCTCAGGAATGTCTGAGAGGAATGACCTTTCTTCTG 1123
Qy 1120 AAGTGTTCACCAAGCCATGCTGTGATGTAATGAGGAGGCACTGAAGCAGCGCTGGCA 1179

Db 1124 AGGTGTTCCATCAAGCCAGCGTGGATGTCACCGAGGAGGCACTGTGCGAGCTGTTGGGA 1183
Qy 1180 CAGGAGGTGTTATGACAGGGAGAACTCGGACATGAGAGGCCACAGTTTGTGGCAGATCATC 1239
Db 1184 CTGGGCGAGTTATGACAGGGAAGAACTGGCCATGGTGGCCACAGTTTGTGGCCGATCATC 1243
Qy 1240 CGTTTCTTTTCTTATTTATGACATAAGATAACCAAGTGCATTTTATTTTTCGSCAGATTT 1299
Db 1244 CTTTCTTTTCTTATCATGACAAATTTACCCACACGATACTATTTGTTGGTAGATCT 1303
Qy 1300 GTCACCCCTAAACTAAGCGTGTCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTC 1359
Db 1304 CTTCACTCCCTAAAGGGAAGACCTATTTCCACATGAGGTTTGTAGCATGAATTAAGC 1363
Qy 1360 CTCAGAACTGCTATTTCAAAATTCGCCAAAAATTTAGAGATGTTTCTACATATTTCTGCTC 1419
Db 1364 CTCAGAACTGCACTCTCAAGTGCCTAAAGTTTAAATACTTCTTACACATTTTAT --- 1419
Qy 1420 TTCTGCAACACTCTCTGCTACCCACTAAATAAAGC--ACAGAAATAATTAGACAATTTGTC 1477
Db 1420 -----ACTTCTGCTATACACTAAATAATACTAAAGCAATTTGTATACGACTTTC 1470
Qy 1478 TATTAATAACATGACAAACCTTATTAATCAATTTGGTCTTTCTAAAAATGGGATCATGCCATTT 1537
Db 1471 AGTGCCTTACAGTATAACTCTATTAATGATTTGTTCTTAAAGTCAGATGATGTCTATTT 1530
Qy 1538 AGATTTTCTCTACTATCAGTTTATTTTATAACATTAACCTTTTACTTTGTTATTTATTTAT 1597
Db 1531 AGTTCATCTCTTATCTGTTTCTTTTAACTTTAGTTTTCAGTGTATTTATTT--G 1589
Qy 1598 TTTATATAATGGTGTGAGTTTTTAAATTTATGCTCACTGCTCTATTTAAATGTAGCTAATAAG 1657
Db 1590 TTTATATAATGGTGTGTTTTTAAATTTGTTGCTCTGTTTAAAGAACT-----GTAACA 1644
Qy 1658 TTATGAGCAGATGATCTGTTTAAATTTCTCTATCTTAATAAATGCTTTTAAATGTTCTCAT 1717
Db 1645 CTACAGAGCAGAAATTTAGATAATTTCTTATTTAAGAAATTCAGCAATTAATTAAT 1704
Qy 1718 ATGAAGAAATAGTAGGTATCCCTCCATGCTCTCTGTAATAAATCTGGAATAAACATTT 1777
Db 1705 AATGAAGGAAAAATATGAGTCTTCCATCTTCCATGATTTCCACCAAGAAAAATGTA 1764
Qy 1778 AAACATAGGCAATATATGTTATGTCATTTCTAGAAATCTAGAAATACATACATATATATGTC 1837
Db 1765 CTTAACAAGACATGT-----TATATCTCTATCTATTTATATATCATATGATGCTATC 1816
Qy 1838 TGTATCTTATTTCAATTTGCAAGTATATAATAATAAACCTGCTTTCCAAACA 1889
Db 1817 TGCACTCTCTATTAATTAGGACTACATCATATAAGTAAGCATGCTTACTTACA 1868

RESULT 6

AK090049
LOCUS

DEFINITION

1952 bp mRNA linear HTC 20-SEP-2003
Mus musculus blood RCB-0035 WEHI-3 cDNA, RIKEN full-length enriched
library, clone:G430078L13 product:serine (or cysteine) proteinase
inhibitor, clade B (ovalbumin), member 2, full insert sequence.

ACCESSION

AK090049

VERSION

AK090049.1 GI:26354876

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus

ORGANISM

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

1 Carninci, P. and Hayashizaki, Y.

AUTHORS

TITLE

High-efficiency full-length cDNA cloning

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

99273253

PUBMED

10349636

REFERENCE

AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishii, K., Kikunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
Riken integrated sequence analysis (RISA) system-384-format
Genome Res. 10 (11), 1757-1771 (2000)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-699 (2001)

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

6 (bases 1 to 1952)
Adachi, J., Aizawa, K., Akimura, T., Arawaka, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurimura, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp)
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES
source

1. .1952
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="FANTOM DB:G430078L13"
/db_xref="MGI:2428955"
/db_xref="taxon:10090"
/clone="G430078L13"
/cell_line="RCB-0035 WEHI-3"
/issue_type="blood"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
54. .1301
/note="unnamed protein product; putative serine (or cysteine) proteinase inhibitor, clade B

CDS

Query Match 52.5%; Score 1000.8; DB 11; Length 1952;
Best Local Similarity 73.0%; Pred. No. 1.2e-176;
Matches 1366; Conservative 0; Mismatches 477; Indels 29; Gaps 5;
ORIGIN

polyA_signal
polyA_site
notes="putative"
notes="putative"

20 ACAGCAACTCAGAGAATAACAGAGAACCAACAGATGAAACAATGGAGGATCTTTGTGT 79
11 ATAGCAATAGTTACTATATTGAGCTGTGTAGAGGATTGAACAATGGAAGACTTTCCAT 70
80 GCGAACAACACTTTTGGCCCTCAATTTTCAAGCATCTGCGAAGAGAGAGCCACCCCA 139
71 GCGAACAACACTTTTGGCCCTCAATTTTCAAGCATGAGAAATCAAACTCTACCCCA 130
140 GAACCTCTTCTCTCCCTCCAGCATCTCGTCCACCATGGCCATGGTCTACATGGGCTC 199
131 GAACATCTTATCTCTCCATGGAGCATCTCATCAACATGGCCATAGTTCTCTCGGTGC 190
200 CAGGGCAGACCGAAGACAGATGGCCAGAGTCTTCAGTTTATGAAGTGGAGCCAA 259
191 TGGGGTAAACACTGAAACAGCAGATGGCCAAAGTCTCGAGTTTATGAATTTGGCAGTTA 250
260 TGCAGTTTACCCCATCACTCCAGAGAACTTATACCAGCTGTGGTTCATGCAGCAGATCCA 319
251 TGGTATCACCACAAGAAACCCAGAGAACTTCAGTGGTGTGATTTCCGACACAGATACA 310
320 GAAGGTAGTTATCTCGATCGGATTTTGGAGGACAGAGCTCCAGATAAAATCCATTCATC 379
311 GAAGGAAATTTATCTCTAGTCTATTTTACAGGCACAGAGGAGATAAAATCCATTCAGC 370
380 CTTCCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGAATTTATTTACTGGAAAGTGT 439
371 CTTCTCTCTCTCAGCTCAACATCAACACACAGGGGATTTTGTAGAAAGTGC 430
440 CAATAAGCTGTTTGGTGAAGTCTCGAGCTTCCGGGAGAGATAATATTGCACTCTGTCA 499
431 AAACAGCTGTTTGGGAGAGAGTCTGCAAGATTCAAGAGAGAGATACATCCAACTCTCAA 490
500 GAAATATTACTCTCAGAACCCAGGAGTAGACTTCTCTAGAAATGTGCAGAGAGAGCTAG 559
491 GAAATATTACTCAACAGACACAGAGCAGTGGACTTCTTGAATGTGTGAAGAGCTAG 550
560 AAAAAAGATTATTTCTGGGTCAAGACTCAACCAAGAGGCAAAATCCCAAACTTTGTACC 619
551 GGAAGAGATTATTTCTGGGTCAAGACTCAACCAAGAGGTAATCCCAAACTCTCTACC 610
620 TGAAGTCTCTAGATGGGATACAGGATGGTCTCGTGAATGTCTGTCTACTTCAAGG 679
611 CGAAGTCTCTAGATGAAGACACAGATGGTGTGATGTCTGTCTACTTCAAGG 670
680 AAGTGGAAATCTCCATTTGAGAGAGAACTAAATGGGCTTTATCCTTTCCGTGTAACCTC 739
671 AAGTGGAAATCTCCATTTGAGAGAGAACTAAATGGGCTTTATCCTTTCCGTGTAACCTC 730
740 GGTCTAGCCACACCTGTACAGATGATGTCTCGTGAAGAGCTAAACATTTGGATACAT 799

731 GCATGAGCATACCTGTCAGATGATGTTCTCCATGCAAGCTGAACATTGATACAT 790
 800 AGAAGACCTAAAGGCTCAGATCTTAGAAGTCCCATATGCTCGAGATGTTAGCATGTTCT 859
 791 AAAGGACCTGAAGACTCAGATCTTAGAAGTCCCATATGCTCGAGATGTTAGCATGTTCT 850
 860 GTTGCTTCAGATGAATTCGCGATGTCACATGCTGCTGAGCTGCTGGAAGTGAAT 919
 851 GTTGCTTCGCGATGAGATGAGGAGCATCCTGCTGGAATGCTGGAAGTGAAT 910
 920 AACCTATGACAACTCAACAGTGAACAGCAAGCAAAATGCTGGAAGTGAAT 979
 911 AACCTTGCCAACTCAACAGTGAACAGCAAGCAAAATGCTGGAAGTGAAT 970
 980 GGTATACATACCCAGTTCAAATTTAGAGGACATTTAGCAATTCAGATCCATTTCTGAAAG 1039
 971 GGTATACATTCCTCAAGTTCAAATTTAGAGGACATTTAGCAATTCAGATCCATTTCTGAAAG 1030
 1040 CATGGCATGAGGAGACCCCTTCAACAGAGGACGGGCAATTTCTCAGGATGTCGAGAG 1099
 1031 CATGGCATGAGGAGACCCCTTCAACAGAGGACGGGCAATTTCTCAGGATGTCGAGAG 1090
 1100 GAATGACCTGTTCTTCTGAGTGTTCACCAAGCCATGCTGAGTGAATGAGGAGG 1159
 1091 GAATGACCTTCTTCTGAGTGTTCACCAAGCCATGCTGAGTGAATGAGGAGG 1150
 1160 CATGAGGACGCGCTGCGACAGAGGATGTTATGACAGGAGAACTGACATGAGGAGG 1219
 1151 CATGAGGACGCGCTGCGACAGAGGATGTTATGACAGGAGAACTGACATGAGGAGG 1210
 1220 ACAGTTGTTGGCAGATCATCGCTTCTTCTTATGATGATGATGATGATGATGATGAT 1279
 1211 ACAGTTGTTGGCAGATCATCGCTTCTTCTTATGATGATGATGATGATGATGATGAT 1270
 1280 TTATTTTTCGAGATTTGCTCAACCTAAACCTAAGCGTGTCTCTGCAAAAGATT 1339
 1271 ACTATTTGTTGATGATTTCTCTCAACCTAAAGGAGGAGGAGGAGGAGGAGGAGG 1330
 1340 TTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1399
 1331 TTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1390
 1400 TTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1457
 1391 TTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1437
 1458 GAATTAATGACAAATGCTTATTAATGACAAATGCTTATTAATGACAAATGCTTATTA 1517
 1438 GAATTAATGACAAATGCTTATTAATGACAAATGCTTATTAATGACAAATGCTTATTA 1497
 1518 AAATGGATCAAGCCATTTAGATTTCTTCTTATTAATGACAAATGCTTATTAATGACAA 1577
 1498 AAATGGATCAAGCCATTTAGATTTCTTCTTATTAATGACAAATGCTTATTAATGACAA 1557
 1578 AAATGGATCAAGCCATTTAGATTTCTTCTTATTAATGACAAATGCTTATTAATGACAA 1637
 1558 AAATGGATCAAGCCATTTAGATTTCTTCTTATTAATGACAAATGCTTATTAATGACAA 1616
 1638 AAATGGATCAAGCCATTTAGATTTCTTCTTATTAATGACAAATGCTTATTAATGACAA 1697
 1617 AAATGGATCAAGCCATTTAGATTTCTTCTTATTAATGACAAATGCTTATTAATGACAA 1671
 1698 AAATGGATCAAGCCATTTAGATTTCTTCTTATTAATGACAAATGCTTATTAATGACAA 1757
 1672 AAATGGATCAAGCCATTTAGATTTCTTCTTATTAATGACAAATGCTTATTAATGACAA 1731
 1758 AAATGGATCAAGCCATTTAGATTTCTTCTTATTAATGACAAATGCTTATTAATGACAA 1817
 1732 AAATGGATCAAGCCATTTAGATTTCTTCTTATTAATGACAAATGCTTATTAATGACAA 1783
 1818 AAATGGATCAAGCCATTTAGATTTCTTCTTATTAATGACAAATGCTTATTAATGACAA 1877
 1784 AAATGGATCAAGCCATTTAGATTTCTTCTTATTAATGACAAATGCTTATTAATGACAA 1843

QY 1878 TGCTTCCAAACA 1889
 DB 1844 TGCTTACTTACA 1855

RESULT 7
 AL579236/c

LOCUS
 DEFINITION

AL579236 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
 clone CS0DH002XJ05 3-PRIME, mRNA sequence.

AL579236
 AL579236.2 GI:31317520

EST.
 SOURCE

Homo sapiens (human)
 ORGANISM

REFERENCE
 AUTHORS

TITLE
 JOURNAL

COMMENT
 On Feb 16, 2001 this sequence version replaced gi:12944080.

Contact: Genoscope
 Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster

9202.f For more information about this cluster, see
 http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DH002CE03NP1&cluster=9202.f. Contact :
 Peng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DH002CE03NP1.

FEATURES
 Location/Qualifiers

1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DH002XJ05"
 /tissue_type="T CELLS (JURKAT CELL LINE)"
 /cell_line="JURKAT CELL LINE"
 /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
 /notes="vector: PCWSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN

Query Match 50.9%; Score 970.4; DB 9; Length 1201;
 Best Local Similarity 88.7%; Pred. No. 6.3e-171;
 Matches 989; Conservative 77; Mismatches 46; Indels 3; Gaps 3;

QY 747 CGCACCTCTACAGATGATGATGCTGGTGAAGCTAAACATTCGATACATAGAAC 806
 DB 1113 CGSCMCCCTTAAVGTGTTGTTGGGAAAGAAAGTGAAGTGAAGTGAAGTGAAGTGA 1054

QY 807 CTAAAGGCTCAGATT-CTAGAACTCCCATATGCTGGAGATGTTAGCATGTTCTTTGCT 865
 DB 1053 CTAAAGGCTCAGATT-CTAGAACTCCCATATGCTGGAGATGTTAGCATGTTCTTTGCT 994

QY 866 TCAGATGAATTCGCGATGTCACCTGGCTGGAGCTGCTGGAAGTGAAGTGAAGTGAAGTGA 925
 DB 993 CCCATATGAATTCGCGATGTCACCTGGCTGGAGCTGCTGGAAGTGAAGTGAAGTGAAGTGA 934

QY 926 TGACAACTCAACAGTGGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 985
 DB 933 TGACAACTCAACAGTGGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 875

QY 986 CATACCCAGTTCAAATTTAGAGGAGCATTTATGAATCCAGATCCATTCCTGAAAGCATGGG 1045

1471 AATTGCTATTATTAACATGACACCCCTATTATTAATCATTTGGTCTCTTAATAATGGATCATG 1530
 390 AATTGCTATTATTAACATGACACCCCTATTATTAATCATTTGGTCTCTTAATAATGGATCATG 331
 1531 CCCATTAGATTTTCCCTTACTATCAGTTATTTTATPAACATTAACCTTTTACTTTGTAT 1590
 330 YCAATTTAGATTTTCCCTTACTATCAGTTATTTTATPAACATTAACCTTTTACTTTGTAT 271
 1591 TTATTATTATTATATATGTTAGTTTAAATATTGCTCAGTCCCTATTAAATGTAGCT 1650
 270 TTATTATKTTATATATGTTAGTTTAAATATTGCTCAGTCCCTATTAAATGTAGCT 211
 1651 AATAAGTTATAGAACGAGATGCTGTTAAATTTCCCTATCTATATAATGCTTTAAATGT 1710
 210 AATAAGTTATAGAACGAGATGCTGTTAAATTTCCCTATCTATATAATGCTTTAAATGT 151
 1711 TCTCATTAATGAAGAATAAGTAGGTATCCCTCCATGCTCTCTGTTAAATAATATCTGGAAA 1770
 150 TCTCATTAATGAAGAATAAGTAGGTATCCCTCCATGCTCTCTGTTAAATAATATCTGGAAA 91
 1771 AATCATTAACATAGCAATATATGTTATGTCATTTCTAGAAATACATAACACATAT 1830
 90 AATCATTAACATAGCAATATATGTTATGTCATTTCTAGAAATACATAACACATAT 31
 1831 ATATGCTGTATCTTATTAATCAATGCAAG 1860
 30 ATATGCTGTATCTTATTAATCAATGCAAG 1

RESULT 9
 AL552852 1044 bp mRNA linear EST 31-MAY-2003
 LOCUS AL552852 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CSODI068YK02 5-PRIME, mRNA sequence.
 ACCESSION AL552852
 VERSION AL552852.2 GI:31274667
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1044)
 Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12892133.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9202.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSODI068BF01Q1&cluster=9202.f. Contact :
 Feng Liang Email: fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSODI068BF01Q1.
 Location/Qualifiers
 1. 1044
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI068YK02"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /notes="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end cloned, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
 source

Query Match 48.9%; Score 932.6; DB 9; Length 1044;

ORIGIN

Best Local Similarity 97.9%; Pred. No. 7.3e-164;
 Matches 962; Conservative 4; Mismatches 13; Indels 4; Gaps 2;
 QY 28 TCAGAGATTAACAGAGAACCAACAGATTGAACAATGGAGGATCTTTGTGTGCAACA 87
 DB 63 TCAGAGATTAACAGAGAACCAACAGATTGAACAATGGAGGATCTTTGTGTGCAACA 122
 QY 88 CACTCTTTGGCCCTCAATTTATTCAAGCATCTGCAAAAGCAAGCCCAACCCAGAACCTCT 147
 DB 123 CACTCTTTGGCCCTCAATTTATTCAAGCATCTGCAAAAGCAAGCCCAACCCAGAACCTCT 182
 QY 148 TCCTCTCCCATGGAGCATCTGTCACCATGCGCATGGTCTACATGGGCTCCAGGGGCA 207
 DB 183 TCCTCTCCCATGGAGCATCTGTCACCATGCGCATGGTCTACATGGGCTCCAGGGGCA 242
 QY 208 GCACGAGACACAGATGGCCAA---GGTGTTCAGTTTAAATGAATGGAGGCAATGCG 264
 DB 243 GCACGAGAAAAAAGAAAAAAGAGAAAMWATTTTAAATGAATGGAGGCAATGCG 302
 QY 265 TTACCCCTCATGCTCCAGAGAACTTTTACCAAGCTGTGGTTTCATCCAGCAGATCCAGAAGG 324
 DB 303 TTACCCCTCATGCTCCAGAGAACTTTTACCAAGCTGTGGTTTCATCCAGCAGATCCAGAAGG 362
 QY 325 GTAGTTATCTGATGCGATTTTTCAGGCAAGCTGCAGATAAATCCATTTCCTTCC 384
 DB 363 GTAGTTATCTGATGCGATTTTTCAGGCAAGCTGCAGATAAATCCATTTCCTTCC 422
 QY 385 GCTCTCAGCTCTCAATCAATGCAATCCAGGGAATTTTACTGGAAGATGTCAATA 444
 DB 423 GCTCTCAGCTCTCAATCAATGCAATCCAGGGAATTTTACTGGAAGATGTCAATA 482
 QY 445 AGCTGTTTGGTGAGAAAGTCTGCGAGCTTCCGGGAAGAAATATATTCGACTCTGTCAAGAT 504
 DB 483 AGCTGTTTGGTGAGAAAGTCTGCGAGCTTCCGGGAAGAAATATATTCGACTCTGTCAAGAT 542
 QY 505 ATTACTCTCAGAACCCAGGAGTAGACTTCTTAGAATGTCAGAGAGAGCTAGAGAA 564
 DB 543 ATTACTCTCAGAACCCAGGAGTAGACTTCTTAGAATGTCAGAGAGAGCTAGAGAA 602
 QY 565 AGATTAAATCTCTGGTGCAAGACTCAAAACCAAGGCAAAATCCCAAACTTTTACCTGAAG 624
 DB 603 AGATTAAATCTCTGGTGCAAGACTCAAAACCAAGGCAAAATCCCAAACTTTTACCTGAAG 662
 QY 625 GTTCTGTAGATGGGATACCAGGATGCTCTGGTGAATGCTGTCTACTTCAAGGAAAGT 684
 DB 663 GTTCTGTAGATGGGATACCAGGATGCTCTGGTGAATGCTGTCTACTTCAAGGAAAGT 722
 QY 685 GGAATCTCATTTGAGAGAAACTAAATGGGCTTTATCTCTTCGCTGTAACTCGGCTC 744
 DB 723 GGAATCTCATTTGAGAGAAACTAAATGGGCTTTATCTCTTCGCTGTAACTCGGCTC 782
 QY 745 AGCGCACACCTGTACAGATGATGTACTTGGTGAAAAGCTAAACATTGGATACATAGAAG 804
 DB 783 AGCGCACACCTGTACAGATGATGTACTTGGTGAAAAGCTAAACATTGGATACATAGAAG 842
 QY 805 ACCTAAGGCTCAGATTCTAGACTCCCATATGCTGGAGATGTTAGCATGTTCTTTGTTGC 864
 DB 843 ACCTAAGGCTCAGATTCTAGACTCCCATATGCTGGAGATGTTAGCATGTTCTTTGTTGC 902
 QY 865 TTCAGATGAATTTGCGGATGTGTCCTCTGCTGGAGCTGCTGGAAGTGAATTAACCT 924
 DB 903 TTCAGATGAATTTGCGGATGTGTCCTCTGCTGGAGCTGCTGGAAGTGAATTAACCT 962
 QY 925 ATGACAACTCAACAGTGGACCAAGCAAAAGCAAAATGGCTGAAGATGAAGTTGAGGTAT 984
 DB 963 ATGACAACTCAACAGTGGACCAAGCAAAAGCAAAATGGCTGAAGATGAAGTTGAGGTAT 1022
 QY 985 ACATACCCAGTTCAATTAAGAA 1007
 DB 1023 ACATA-CCCAGTTCAAAATTAGAA 1044

RESULT 10

```

AL576638/c
LOCUS      AL576638      991 bp      mRNA      linear      EST 01-JUN-2003
DEFINITION  AL576638 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0D1078YH02 3-PRIME, mRNA sequence.
ACCESSION  AL576638
VERSION    AL576638.2 GI:31314920
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 991)
            Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
            On Feb 16, 2001 this sequence version replaced gi:12938977.
CONTACT    Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 9202.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/078DD01NP1kcluster-9202.f. Contact :
            Cgi-bin/cluster.cgi?seq=CS0D1078DD01NP1kcluster-9202.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0D1078DD01NP1.
FEATURES   Location/Qualifiers
            1..991
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0D1078YH02"
            /tissue_type="PLACENTA COT 25-NORMALIZED"
            /note="1st strand cDNA was primed with a NotI-oligo(dT)
            primer. Five prime end enriched, double-strand cDNA was
            digested with Not I and cloned into the Not I and EcoR V
            sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 48.8%; Score 931; DB 9; Length 991;
Best Local Similarity 96.7%; Pred.No.1.5e-163; Indels 2; Gaps 2;
Matches 959; Conservative 8; Mismatches 23;
2y 884 TGTGTCCACTGGCTTGGAGCTGCTGGAAGTGAATTAACCTATGACAACTCAACAAAGTG 943
Db 991 TGTGTCCACTGGCTTGGAGCTGCTGGAAGTGAATTAACCTATGACAACTCAACAAAGTG 932
2y 944 GACCAGCAAGACAAAATGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATT 1003
Db 931 GACCAGCAAGACAAAATGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATT 872
2y 1004 AGAAGAGCATTTAAGTCACTCAGATCCATTTCTGAAAGCATGGGATGGAGGACGCTTCAA 1063
Db 871 AGAAGAGCATTTAAGTCACTCAGATCCATTTCTGAAAGCATGGGATGGAGGACGCTTCAA 812
2y 1064 CAAGGACGGGCAATTTCTCAGGAGTTCGGAGAGGAATGACCTGTTCTTCTGGAAGT 1123
Db 811 CAAGGACGGGCAATTTCTCAGGAGTTCGGAGAGGAATGACCTGTTCTTCTGGAAGT 752
2y 1124 GTTCCACAGCCATGGTGGATGTGATGAGGAGGCGACTGACGACGCGCTGCACAGG 1193
Db 751 -KTCACCAAGCCATGGTGGATGTGATGAGGAGGCGACTGACGACGCGCTGCACAGG 693
2y 1194 AGGTGTTATGACAGGAGAACTGGACATGGAGGCCACACAGTTGTGGCAGATCATCCGTT 1243
Db 692 AGGTGTTATGACAGGAGAACTGGACATGGAGGCCACACAGTTGTGGCAGANCAACCCCTT 633
2y 1244 TCTTTTCTTATTATGATACATACCAAGTGCATTTTATTTTTCGGCAGATTTTGCCTC 1303
Db 632 TCTTTTCTTATTATGATACATACCAAGTGCATTTTATTTTTCGGCAGATTTTCTC 573

```

```

Qy 1304 ACCCTAAACAAAGCGTCTGCTTCTGCAAAAGATTTTCTAGATGAGCTGTGCGCTCA 1363
Db 572 ACCCTAAACAAAGCGTCTGCTTCTGCAAAAGATTTTCTAGATGAGCTGTGCGCTCA 513
Qy 1364 GAATTGCTATTTCAAATTCGCCAAAATTTAGAGATGTTTTCTACATATTTCTGCTCTCT 1423
Db 512 GAATTGCTATTTCAAATTCGCCAAAATTTAGAGATGTTTTCTACATATTTCTGCTCTCT 453
Qy 1424 GAACAACTTCTGCTACCCCACTAAATAAAAAACAGAGAAATAATTAGACAAATTTCTATTAT 1483
Db 452 GAACAACTTCTGCTACCCCACTAAATAAAAAACAGAGAAATAATTAGACAAATTTCTATTAT 393
Qy 1484 PACATGACAAACCTATTATATCATTTTGGTCTCTTAAATGGGATCATGCCCATTTAGATT 1543
Db 392 AACATGACAAACCTATTATATCATTTTGGTCTCTTAAATGGGATCATGCCCATTTAGATT 333
Qy 1544 TCCTTACTATCAGTTTATTTTATAAATTAACATTAATTTTACTTTTGTATTATTTATATA 1603
Db 332 CCCTTACTATCAGTTTATTTTATAAATTAACATTAATTTTACTTTTGTATTATTTATATA 273
Qy 1604 TAATGGTGAGTTTTTAAATTTATTTGCTCCTGCTTAAATTAATGATAGCTAATAAAGTTAG 1663
Db 272 TAATGGTGAGTTTTTAAATTTATTTGCTCCTGCTTAAATTAATGATAGCTAATAAAGTTAG 213
Qy 1664 AAGCAGATGATCTGTTAAATTTCTTATCTAATAAATGCTTTAATTTCTCATATGAAG 1723
Db 212 AAGCAGATGATCTGTTAAATTTCTTATCTAATAAATGCTTTAATTTCTCATATGAAG 153
Qy 1724 AATAAGTAGGTATCCCTCCATGCTTCTGTAATAAATATCTGGAAAAAACAATTAAACAA 1783
Db 152 AATAAGTAGGTATCCCTCCATGCTTCTGTAATAAATATCTGGAAAAAACAATTAAACAA 93
Qy 1784 TAGGCARATATATGTTATGTCATTTCTAGAAATACATACATATATATGTC-TGAT 1842
Db 92 TAGGCARATATATGTTATGTCATTTCTAGAAATACATACATATATATGTCNNAT 33
Qy 1843 CTTATATCAATTGCAAGTATATATAATAATPAA 1874
Db 32 CTTATATCAATTGCAAGAAAMNNVGMVMA 1

```

RESULT 11
LOCUS AL553610
DEFINITION AL553610 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1078YH02 5-PRIME, mRNA sequence.
ACCESSION AL553610
VERSION AL553610.2 GI:31275424
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1000)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12893604.
CONTACT Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9202.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/078DD01NP1kcluster-9202.f. Contact :
 Cgi-bin/cluster.cgi?seq=CS0D1078DD01NP1kcluster-9202.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D1078DD01NP1.
FEATURES Location/Qualifiers
 1..1000

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1013Y022"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 48.5%; Score 925.4; DB 9; Length 1000;
Best Local Similarity 99.8%; Pred. No. 1.6e-162;
Matches 937; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GAGGACATTGCCCGTCAGACGCAACTCAGAGAAATTAACAGAGAAACCAACAGATTGAAA 60
DB |
QY 62 GATGAGCATCTTGTGTGGCAACACACTCTTGGCCCTCAATTTATTCAAGCATCTGG 120
DB |
QY 61 CAATGAGGATCTTGTGTGGCAACACACTCTTGGCCCTCAATTTATTCAAGCATCTGG 120
DB |
QY 122 CAATGAGGATCTTGTGTGGCAACACACTCTTGGCCCTCAATTTATTCAAGCATCTGG 181
DB |
QY 121 CAAAAGCAAGCCCAACCCAGAACTCTTCTCTCCCATGAGGATCTCGTCCACCATGG 180
DB |
QY 182 CAAAAGCAAGCCCAACCCAGAACTCTTCTCTCCCATGAGGATCTCGTCCACCATGG 241
DB |
QY 181 CCATGCTCTACATGGGCTCCAGGGGAGCAGCCGAGACAGATGCCAAGGTGCTTCAGT 240
DB |
QY 242 CCATGCTCTACATGGGCTCCAGGGGAGCAGCCGAGACAGATGCCAAGGTGCTTCAGT 301
DB |
QY 241 TTAATGAAGTGGAGCCCAATGACGATTAACCCCATGACTCTCCAGAGACATTTACCAGCTTG 300
DB |
QY 302 TTAATGAAGTGGAGCCCAATGACGATTAACCCCATGACTCTCCAGAGACATTTACCAGCTTG 361
DB |
QY 301 GGTTCATGACGAGATCCAGAAAGGTAGTTATCTGTGATGCGATTTTGAGGCACAAAGCTG 360
DB |
QY 362 GGTTCATGACGAGATCCAGAAAGGTAGTTATCTGTGATGCGATTTTGAGGCACAAAGCTG 421
DB |
QY 361 CAGATAAATCCCATCTCTCCGCTCTCCAGCTCTGCACTCTGCAATCAATCAATCCACAGGGA 420
DB |
QY 422 CAGATAAATCCCATCTCTCCGCTCTCTGCTCTCTGCTCTGCAATCAATCAATCCACAGGGA 481
DB |
QY 421 ATTATTTACTGGAAGTGTCAATGAAGTGTGTTGTGAGAAAGTGTGCGAGCTTCCGGGAAG 480
DB |
QY 482 ATTATTTACTGGAAGTGTCAATGAAGTGTGTTGTGAGAAAGTGTGCGAGCTTCCGGGAAG 541
DB |
QY 481 AATATATTGCACTCTGTCAGAAATATTACTCTCTCAGAACCCCGAGGAGTAGACTTCCCTAG 540
DB |
QY 542 AATATATTGCACTCTGTCAGAAATATTACTCTCTCAGAACCCCGAGGAGTAGACTTCCCTAG 601
DB |
QY 541 AATGTGAGAAAGAGCTAGAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCA 600
DB |
QY 602 AATGTGAGAAAGAGCTAGAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCA 661
DB |
QY 601 AATATCCCAAACTGTACTCTGAAGGTTCTGTAGATGGGGATACCAAGGATGGTCTTGGTGA 660
DB |
QY 662 AATATCCCAAACTGTACTCTGAAGGTTCTGTAGATGGGGATACCAAGGATGGTCTTGGTGA 721
DB |
QY 661 ATGCTGTCTACTTCAAGGAAAGTGGAAATCTCAATTTGAGAGAACTTAATATGGGCTTTT 720
DB |
QY 722 ATGCTGTCTACTTCAAGGAAAGTGGAAATCTCAATTTGAGAGAACTTAATATGGGCTTTT 781
DB |
QY 721 ATCCTTTCCGTGTAAACTCGGCTCAGCCGACACCTGTGACAGATGATGTAATCTTGGTGA 780
DB |
QY 782 ATCCTTTCCGTGTAAACTCGGCTCAGCCGACACCTGTGACAGATGATGTAATCTTGGTGA 841
DB |
QY 781 AGCTTAACATTTGGATACATAGAGACCTTAAGGCTCAGATTCTAGAACTCCCATATGCTG 840
DB |
QY 842 AGCTTAACATTTGGATACATAGAGACCTTAAGGCTCAGATTCTAGAACTCCCATATGCTG 901
DB |
QY 841 GAGATGTAGCAT-GTTCTCTGTGCTTCCAGATGAATAATGGCCGATGTGTCACCTGCTTG 899
```

```

DB 902 GAGATGTAGCATGTTCTTGTGTTGCTTCCAGATGAATTCGCGATGTGTCCTGCTG 961
QY 900 GAGCTGCTGGAAGTGAATAATACCTATGACAAACTCAAC 938
DB 962 GAGCTGCTGGAAGTGAATAATACCTATGACAAACTCAAC 1000

RESULT 12
BX396200
LOCUS BX396200
DEFINITION BX396200 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
ACCESSION clone CS0D1013Y022 5-PRIME, mRNA sequence.
VERSION BX396200
KEYWORDS BX396200.1 GI:30628769
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1041)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9202.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1013BH1Q1&cluster=9202.f. Contact :
Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0D1013BH1Q1.

FEATURES
Location/Qualifiers
1..1041
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1013Y022"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 47.9%; Score 913; DB 13; Length 1041;
Best Local Similarity 91.1%; Pred. No. 3.3e-160;
Matches 894; Conservative 68; Mismatches 18; Indels 1; Gaps 1;

QY 5 AGCATTCCTCGTCAGACGCAACTCAGAGAAATTAACAGAGAAACCAACAGATTGAAACAAT 64
DB 61 AGCATTCCTCGTCAGACGCAACTCAGAGAAATTAACAGAGAAACCAACAGATTGAAACAAT 120
QY 65 GGAGGATCTTTGTGTGGCAACACACTCTTTCCTCTCAATTTATTCAAGCATCTGCGAAA 124
DB 121 GGAGGATCTTTGTGTGGCAACACACTCTTTCCTCTCAATTTATTCAAGCATCTGCGAAA 180
QY 125 AGCAAGCCCAACCCAGAACTCTTCTCTCCCATGGAGCATCTCGTCCCAATGGCCAT 184
DB 181 AGCAAGCCCAACCCAGAACTCTTCTCTCCCATGGAGCATCTCGTCCCAATGGCCAT 240
QY 185 GGTCTCATGGGCTCCAGGGGAGCAGCAGCAGAGACAGATGGCAGAGTGTGCTTCAAGTTAA 244
DB 241 GGTCTCATGGGCTCCAGGGGAGCAGCAGCAGAGACAGATGGCAGAGTGTGCTTCAAGTTAA 300
QY 245 TGAAGTGGGAGCCCAATGCAGTTACCCCATGACTCCAGAGAACTTTACCAGCTGTGGGTT 304
DB 301 TGAAGTGGGAGCCCAATGCAGTTACCCCATGACTCCAGAGAACTTTACCAGCTGTGGGTT 360
```

```
305 CATGCACGATCCAGAGGGTAGTTATCCGATCGGATTTTCAGGCAACAGCTGAGA 364
361 CATGCACGATCCAGAGGGTAGTTATCCGATCGGATTTTCAGGCAACAGCTGAGA 420
365 TAAATCAATTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAATTA 424
421 TAAATCAATTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAATTA 480
425 TTTACTGGAAAGTGTCAATAAGCTGTTGGTGAGAAAGTCTCGAGGCTTCGGGGAAGATA 484
481 TTTACTGGAAAGTGTCAATAAGCTGTTGGTGAGAAAGTCTCGAGGCTTCGGGGAAGATA 540
485 TATTCGACTCTGTGAGAAATATTAATCTCTCAGAACCCAGGAGTAGACTTCTAGAAATG 544
541 TATTCGACTCTGTGAGAAATATTAATCTCTCAGAACCCAGGAGTAGACTTCTAGAAATG 600
545 TGCAGAGAAAGCTAGAAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAAT 604
601 TGCAGAGAAAGCTAGAAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAAT 660
605 CCNAACTGTACTGTAAGGTTCTGTAGATGGGATACAGGATGCTCTGTGTAATGC 664
661 CCNAACTGTACTGTAAGGTTCTGTAGATGGGATACAGGATGCTCTGTGTAATGC 720
665 TGTCTACTTCAAAGGAAAGTGGAAAGTCAATCTTTCAGAGAAAGTCAATCTTATCC 724
721 TGTCTACTTCAAAGGAAAGTGGAAAGTCAATCTTTCAGAGAAAGTCAATCTTATCC 780
725 TTTCCGCTGTAAGTGGGCTCAGGCAACCTCTACAGATGATGATCTGGTGAAGCT 784
781 TTTCCGCTGTAAGTGGGCTCAGGCAACCTCTACAGATGATGATCTGGTGAAGCT 840
785 AAACATTTGATACATAGAGACCTAAAGCTCAGATTTCTAGAACTCCCATATCTGGAGA 844
841 AAACATTTGATACATAGAGACCTAAAGCTCAGATTTCTAGAACTCCCATATCTGGAGA 900
845 TGTAGCATGTTCTGTTGCTTCCAGATGATGATGATGATGATGATGATGATGATGATGAT 904
901 TGTAGCATGTTCTGTTGCTTCCAGATGATGATGATGATGATGATGATGATGATGATGAT 960
905 GCTGGAAAGTCAATATGCAAACTC-AACAAGTGGACCAAGCAAGCAAAATGG 963
961 GCTGGAAAGTCAATATGCAAACTC-AACAAGTGGACCAAGCAAGCAAAATGG 1020
964 CTGAAGATGAAGTTGAGGTAT 984
1021 CTGAGTGAAGTTGGGTTT 1041
```

```
RESULT 13
AL557133
LOCUS
DEFINITION
AL557133 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
clone CS0DH002YJ05 5-PRIME, mRNA sequence.
ACCESSION
AL557133
VERSION
AL557133.2 GI:31278933
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12900443.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
```

```
9202.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DH002CE03QPl&cluster=9202.f. Contact :
Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DH002CE03QPl.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH002YJ05"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
```

ORIGIN

```
Query Match 46.5%; Score 888; DB 9; Length 1201;
Best Local Similarity 98.1%; Pred. NO. 1.5e-155;
Matches 926; Conservative 4; Mismatches 9; Indels 5; Gaps 3;

Qy 3 GGAGCATTTGCCGCTCAGACAGCAACTCAGAGAAATAACACAGAGAAACACAGATTGAACA 62
Db 48 GGAGCATTTGCCGCTCAGACAGCAACTCAGAGAAATAACACAGAGAAACACAGATTGAACA 107
Qy 63 ATGGAGGATCTTTGTGTGGCAACACACTTTTGGCCCTCAATTTATTCAGCATCTGCA 122
Db 108 ATGGAGGATCTTTGTGTGGCAACACACTTTTGGCCCTCAATTTATTCAGCATCTGCA 167
Qy 123 AAAGCAAGCCCCACCCAGCAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGCC 182
Db 168 AAAGCAAGCCCCACCCAGCAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGCC 227
Qy 183 ATGTCTACATGGGCTCCAGGGGAGCAGCAGAGACAGATGCGCAAGGTCTTCAGTTT 242
Db 228 ATGTCTACATGGGCTCCAGGGGAGCAGCAGAGACAGATGCGCAAGGTCTTCAGTTT 287
Qy 243 AATGAAGTGGAGCCCAATGCACTTACCCCATGACTCCAGAGAACTTTACAGCTGTGG 302
Db 288 AATGAAGTGGAGCCCAATGCACTTACCCCATGACTCCAGAGAACTTTACAGCTGTGG 347
Qy 303 TTCTGACGAGATCCAGAAAGGTAGTTATCTCTGATGCGATTTTGCAGGACAGAGCTGCA 362
Db 348 TTCTGACGAGATCCAGAAAGGTAGTTATCTCTGATGCGATTTTGCAGGACAGAGCTGCA 407
Qy 363 GATAAATCCATTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAAT 422
Db 408 GATAAATCCATTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAAT 467
Qy 423 TATTTACTGGAAAGTGTCAATAAGCTGTTTGGTGAGAAAGTCTCGAGCTTCCGGGAAGAA 482
Db 468 TATTTACTGGAAAGTGTCAATAAGCTGTTTGGTGAGAAAGTCTCGAGCTTCCGGGAAGAA 527
Qy 483 TATTTCTGACTCTGTGAGAAATATTAATCTCTCAGAACCCAGGAGTAGACTTCTAGAA 542
Db 528 TATTTCTGACTCTGTGAGAAATATTAATCTCTCAGAACCCAGGAGTAGACTTCTAGAA 587
Qy 543 TGTGAGAAAGAGCTAGAAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAA 602
Db 588 TGTGAGAAAGAGCTAGAAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAA 647
Qy 603 ATCCCAAACTTTTACCTGAAGGTTCTGTAGATGGGATACAGGATGCTCTGCTGTAAT 662
Db 648 ATCCCAAACTTTTACCTGAAGGTTCTGTAGATGGGATACAGGATGCTCTGCTGTAAT 707
Qy 663 GCTGCTACTTCAAAGGAAAGTGGAAAACTCCATTTTGAAGAAAGCAATTAATGGGCTTAT 722
Db 708 GCTGCTACTTCAAAGGAAAGTGGAAAACTCCATTTTGAAGAAAGCAATTAATGGGCTTAT 767
```

```

QY 723 CTTTCCGTTAACTCGGCTAGCGCACACCTGTACAGATGATGTACTTCGCTGAAAAG 782
Db 768 CTTTCCGTTAACTCGGCTAGCGCACACCTGTACAGATGATGTACTTCGCTGAAAAG 827
QY 783 CTAACATTGGATACATAGAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGGA 842
Db 828 CTAACATTGGATACATAGAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGGA 887
QY 843 GATGTTAGCATGTTCTGTTGCTTCCAGATGAATTCGCGATGTCACACTGGCTTGGAG 902
Db 888 GATGTTAGCATGTTCTGTTGCTTCCAGATGAATTCGCGATGAATTCGCG--TGTGCTACTGGCTTGGAG 945
QY 903 CTGCTGGAAGTGAATAAATCACTATGACAACTCAACAAGTGGAC 946
Db 946 CTGCTGGAAR--TGAATAACTATGAC-AACTCMACAAGTGGCC 986

RESULT 14
LOCUS AL544520 959 bp mRNA linear EST 31-MAY-2003
DEFINITION AL544520 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSOD1021YJ18 5-PRIME, mRNA sequence.
ACCESSION AL544520
VERSION AL544520.2 GI:31266363
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 959)
Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12877000.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9202.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOD1021DE09QP1&cluster=9202.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSOD1021DE09QP1.
FEATURES
Location/Qualifiers
1..959
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1021YJ18"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 46.4%; Score 884.6; DB 9; Length 959;
Best Local Similarity 99.0%; Pred. No. 6,6e-155;
Matches 906; Conservative 4; Mismatches 3; Indels 2; Gaps 2;

QY 10 TCCCGCTCAGACGCACTCAGAGATTAACAGAGAACACCAAGATTGAACAATGGAGG 69
Db 42 TGCCGCTCAGACGCACTCAGAGATTAACAGAGAACACCAAGATTGAACAATGGAGG 101
QY 70 ATCTTTGTGGGAAACACACTCTTTCCTCAATTTATTCACGATCTGGCAAAAGCAA 129
Db 102 ATCTTTGTGGGAAACACACTCTTTCCTCAATTTATTCACGATCTGGCAAAAGCAA 161

```

```

QY 130 GCCCACCAGACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCT 189
Db 162 GCCCACCAGACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCT 221
QY 190 ACATGGGCTCCAGGGCAGCACCGAAGACAGATGGCCAAAGTGCTTTCAGTTTAATGAAG 249
Db 222 ACATGGGCTCCAGGGCAGCACCGAAGACAGATGGCCAAAGTGCTTTCAGTTTAATGAAG 281
QY 250 TGGAGCCCAATGAGTTTACCCCATGACTCCAGAGAACTTTACAGAGCTGTGGTTTCATGC 309
Db 282 TGGAGCCCAATGAGTTTACCCCATGACTCCAGAGAACTTTACAGAGCTGTGGTTTCATGC 341
QY 310 AGCAGATCCAGAGGGTAGTTTATCTGATGCGATTTTGCAGGCAACAGCTGCAGATAAAA 369
Db 342 AGCAGATCCAGAGGGTAGTTTATCTGATGCGATTTTGCAGGCAACAGCTGCAGATAAAA 401
QY 370 TCCATTATCTCTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTATTAC 429
Db 402 TCCATTATCTCTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTATTAC 461
QY 430 TGGAAAGTGTCAATAAGCTGTTTGGTGAGAGTCTGGAGCTTCGGGAAGATATATTC 489
Db 462 TGGAAAGTGTCAATAAGCTGTTTGGTGAGAGTCTGGAGCTTCGGGAAGATATATTC 521
QY 490 GACTCTGTGAGAAATATTTACTCTCAGAACCCCAAGGAGTAGACTTCTCTAGAAATGTGCAG 549
Db 522 GACTCTGTGAGAAATATTTACTCTCAGAACCCCAAGGAGTAGACTTCTCTAGAAATGTGCAG 581
QY 550 AAGAGCTAGAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAA 609
Db 582 AAGAGCTAGAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAA 641
QY 610 ACTTGTACTCTGAAGTCTGTAGATGGGATACAGAGATGGTCTGTGTGAATGTGTCT 669
Db 642 ACTTGTACTCTGAAGTCTGTAGATGGGATACAGAGATGGTCTGTGTGAATGTGTCT 701
QY 670 ACTTCAAGGAGAGTGGAAAACCTCAATTTGAGAGAACTAAATGGGGTTTATCTTTCC 729
Db 702 ACTTCAAGGAGAGTGGAAAACCTCAATTTGAGAGAG-ACTTAATGGGGTTTATCTTTCC 760
QY 730 GTGTAACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGGTGAAGAGCTAAACA 789
Db 761 GTGTAACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGGTGAAGAGCTAAACA 820
QY 790 TTGGATACATAGAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGATGTTA 849
Db 821 TTGGATACATAGAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGATGTTA 880
QY 850 GCATGTTCTTGTGCTTCCAGATGAATTCGCGATGTGTCCACT-GGCTTGGAGCTGCTG 908
Db 881 GCATGTTCTTGTGCTTCCAGATGAATTCGCGATGTGTCCACTGGGCTTGGAGCTGCTG 940
QY 909 GAAAGTGAATAAACC 923
Db 941 GAAAGTGAATCAACY 955

RESULT 15
LOCUS AL545835 989 bp mRNA linear EST 31-MAY-2003
DEFINITION AL545835 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSOD1023Y006 5-PRIME, mRNA sequence.
ACCESSION AL545835
VERSION AL545835.2 GI:31267670
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 989)
Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)

```


COMMENT On Feb 15, 2001 this sequence replaced gi:12878382.

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segreg@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9202.f For
more information about this cluster, see
<http://www.Genoscope.cns.fr/>
<http://cgl-bin/cluster.cgi?seq=CS01023BH03QPl&cluster=9202.f>. Contact :
Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faratday Avenue Genoscope sequence ID : CS0D1023BH03QPl.

```

FEATURES
    source
        Location/Qualifiers
            1..989
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0D1023Y006"
                /tissue_type="PLACENTA COT 25-NORMALIZED"
                /clone_idb="Homo sapiens PLACENTA COT 25-NORMALIZED"
                /note="1st strand cDNA was primed with a NotI-oligo(dT)
                    primer. Five prime end enriched, double-strand cDNA was
                    digested with Not I and cloned into the Not I and EcoR V
                    sites of the pCMVSPORT 6 vector. Library was normalized."

```

)RIGIN

	Query Match	46.3%; Score 883.4; DB 9; Length 989;	
	Best Local Similarity	97.5%; Pred. No. 1.le-154;	
	Matches	890; Conservative	6; Mismatches 17; Indels 0; Gaps 0;
iy	10	TGCCCGTCTAGACAGCAACTCAGAGAATAAACAGAGAAACAACCAGATTGAAACAATGCAGG	69
ib	64	TGGCGGTGAGAGCAACTCAGAGATAASGAGAGAGAGGATTGAAACAATGCAGG	123
iy	70	ATCTTTGTGTGGAAAACACACTCTTTTGCCCTCAATTTATTCAAGCATCTGGCAAAGCAA	129
ib	124	ATCTTTGTGTGSAACACACTCTTTGSCCTCAATTTATTCAAGCATCTGGCAAAGCAA	183
iy	130	GCCCCACCCAGAACCTCTTCCTCTCCCCTATGGAGCATCTCGTCACCATGCCCATTGGTCT	189
ib	184	GCCCCACCCAGAACCTCTTCCTCTCCCCTATGGAGCATCTCGTCACCATGSCATGGTCT	243
iy	190	ACATGGGCTCCAGGGGAGCACCGAAGACAGATGGCGAAGGTGCTTCAGTTTAATGAAG	249
ib	244	AGATGGGCTSCAGGGCAGCACCGAAGACAGATGGCCAAGGTGCTTCAGTTTAATGAAG	303
iy	250	TGGGAGCCAAATGCAGTTACCCCATGACTCCAGAGAACTTTACCAGCTGTGGGTTCATGC	309
ib	304	TGGGAGCCAAATGCAGTTACCCCATGACTCCAGAGAACTTTACCAGCTGTGGGTTCATGC	363
iy	310	AGCAGATCCAGAAAGGTAGTTATCTGATGGCATTTTCGAGGCACAGCTCCAGATAAAA	369
ib	364	AGCAGATCCAGAAAGGTAGTTATCTGATGGCATTTTCGAGGCACAGCTCCAGATAAAA	423
iy	370	TCCATTTCATCTTCGGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAATTTATTAC	429
ib	424	TCCATTTCATCTTCGGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAATTTATTAC	483
iy	430	TGAAAAAGTGCANATAAGCTGTTGGTGAGAAAGTCTGCGAGCTTCGGGAAGAATATATTC	489
ib	484	TGAAAAAGTGCANATAAGCTGTTGGTGAGAAAGTCTGCGAGCTTCGGGAAGAATATATTC	543
iy	490	GACTCTGTGAGAAATA-TACTCTCCAGAACCCACAGGCAAGTAGACTTCCTAGAAATGTGCAG	549
ib	544	GACTCTGTGAGAAATA-TACTCTCCAGAACCCACAGGCAAGTAGACTTCCTAGAAATGTGCAG	603
iy	550	AAGAAGCTGAAAAAAGATTAAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAA	609
ib	604	AAGAAGCTGAAAAAAGATTAAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAA	663
iy	610	ACTTGTATACCTGAAGGTTCTGTAGATGGGATACACAGGATGGTCTCTGTGTGAATGTGCTCT	669

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

DN nucleic - nucleic search, using sw model
Run on: March 10, 2004, 08:38:17 ; Search time 9432.34 Seconds
(without alignments)
17011.249 Million cell updates/sec

Title: US-10-084-817-30
Perfect score: 3702
Sequence: 1 gcctagggaatgacaggcat.....aagggcggtaccgagaac 3702

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.hcg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.hcg.hum.*
31: em.hcg.inv.*
32: em.hcg.other.*
33: em.hcg.mus.*
34: em.hcg.pln.*
35: em.hcg.tod.*
36: em.hcg.nam.*
37: em.hcg.vrt.*
38: em.sy.*
39: em.hgo.hum.*
40: em.hgo.mus.*
41: em.hgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3116.4	84.2	3610	6	AX330212	AX330212 Sequence
2	3116.4	84.2	3610	6	AX336432	AX336432 Sequence
3	3116.4	84.2	3610	6	AX409760	AX409760 Sequence
4	3116.4	84.2	3610	9	HUMCD5M	M83216 Human aorta
5	3116.4	84.2	3610	11	G28707	G28707 SMS3978 Er
6	2222.6	60.0	4091	9	HSM808959	BX648808 Homo sapi
7	1570.2	42.4	2215	6	AR001263	AR001263 Sequence
8	1570.2	42.4	2215	6	E05382	E05382 DNA encodin
9	1570.2	42.4	2215	6	I23421	I23421 Sequence 14
10	1562	42.2	4130	9	BC040354	BC040354 Homo sapi
11	1558	42.1	2198	9	HUMCAL1	D90452 Homo sapien
12	1522	41.1	2375	9	HUMCALD	M64110 Human calde
13	1481.6	40.0	2137	6	AR001262	AR001262 Sequence
14	1481.6	40.0	2137	6	I23420	I23420 Sequence 13
15	1469.4	39.7	2120	9	HUMLCA2	D90453 Homo sapien
16	1455.8	39.3	3342	6	AX202102	AX202102 Sequence
17	1433.6	38.7	2137	6	E05381	E05381 DNA encodin
18	1312.2	35.4	3555	4	RABRSMC	L37206 Oryctolagus
19	1116.2	30.2	4065	4	RABLCA	L37147 Oryctolagus
20	1096.8	29.6	171530	9	AC083870	AC083870 Homo sapi
21	1087.8	29.4	1677	6	AR001261	AR001261 Sequence
22	1087.8	29.4	1677	6	I23419	I23419 Sequence 12
23	1085.6	29.3	141930	9	AC146389	AC146389 Pan trogl
24	999.2	27.0	1599	6	AR001260	AR001260 Sequence
25	999.2	27.0	1599	6	I23418	I23418 Sequence 11
26	993.2	26.8	164345	9	AC145994	AC145994 Pan trogl
27	993	26.8	1896	4	AY154474	AY154474 Bos tauru
28	937.4	25.3	939	6	AR001259	AR001259 Sequence
29	937.4	25.3	939	6	I23417	I23417 Sequence 10
30	925.2	25.0	4108	5	CHKSMCA	J04968 Chicken smo
31	917.4	24.8	5541	10	RN118419	U18419 Rattus norv
32	909	24.6	2564	5	CHKICAL	M28417 Chicken h-c
33	906.4	24.5	2831	10	BC019435	BC019435 Mus muscu
34	889	24.0	3941	9	AF247820	AF247820 Homo sapi
35	839.4	22.7	2637	9	HSM806613	BX538339 Homo sapi
36	691	18.7	705	6	AX885040	AX885040 Sequence
37	691	18.7	705	6	BD024650	BD024650 Sequence
38	689.2	18.6	756	6	BD079603	BD079603 Cancer-as
39	653.4	17.6	2696	6	HSM809239	BX649088 Homo sapi
40	597.6	16.1	706	6	BD079604	BD079604 Cancer-as
41	551.8	14.9	1126	9	BC014035	BC014035 Homo sapi
42	544.8	14.7	548	6	AX885706	AX885706 Sequence
43	544.8	14.7	548	6	BD025316	BD025316 Sequence
44	528.8	14.3	169199	9	AC090497	AC090497 Homo sapi
45	502	13.6	196487	2	AC109173	AC109173 Mus muscu

ALIGNMENTS

RESULT 1
AX330212
LOCUS AX330212 3610 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 721 from Patent WO0194629.
ACCESSION AX330212
VERSION AX330212.1 GI:18103190
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endreass,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature

2197	2y	GAGGAGGAGCAGGAGGAGCCGATCGAAAACTCAGAGGAGGAGGAGAGGAGGAGGCT	2255
1981	3b	GAGGAGGAGCAGGAGGAGGAGCCGATCGAAAACTCAGAGGAGGAGGAGAGGAGGAGGCT	2040
2257	2y	AAAGGAAGAGATTGAAAGCGGAGAGCAGAACTGCTGAGAAAAGGAGAGGAGAGGAGGAGGCT	2316
2041	3b	AAAGGAAGAGATTGAAAGCGGAGAGCAGAACTGCTGAGAAAAGGAGAGGAGAGGAGGAGGCT	2100
2317	2y	AGATGCGCTTGTCAGATGACAGAGAAACCAATCAAGTGTCTTCACTCTTAAGGTTCACTCTCT	2376
2101	3b	AGATGCGCTTGTCAGATGACAGAGAAACCAATCAAGTGTCTTCACTCTTAAGGTTCACTCTCT	2160
2377	2y	CAAGATAGAGAGCGAGCAGAAATTTTGAATAAGTCTGTGAGAAAAGCAGTGGTGTCAA	2436
2161	3b	CAAGATAGAGAGCGAGCAGAAATTTTGAATAAGTCTGTGAGAAAAGCAGTGGTGTCAA	2220
2437	2y	ATCGAGCCCATCAGCAGCAGCAATAGTCTCCAGATTTGACAGCAGACTCGAGCAGATATACCAG	2496
2221	3b	ATCGAGCCCATCAGCAGCAGCAATAGTCTCCAGATTTGACAGCAGACTCGAGCAGATATACCAG	2280
2497	2y	TGCAATTGAGGGGAACAAAACGCAAAACCTACAAGCGGAGCCTCGGATCTTCTGT	2556
2281	3b	TGCAATTGAGGGGAACAAAACGCAAAACCTACAAGCGGAGCCTCGGATCTTCTGT	2340
2557	2y	TCTGCTGAGGTTGTACGCAACATCAAGAGTATGTGGGAGAAAGGAAATGTGTTTCAATC	2616
2341	3b	TCTGCTGAGGTTGTACGCAACATCAAGAGTATGTGGGAGAAAGGAAATGTGTTTCAATC	2400
2617	2y	CCCACTGCAGCAGGACACCAAAATTAAGGAAACTGTGGCTTGAAAGTAGGGGTTTCTAG	2676
2401	3b	CCCACTGCAGCAGGACACCAAAATTAAGGAAACTGTGGCTTGAAAGTAGGGGTTTCTAG	2460
2677	2y	CGGCATCAATGAATGGCTAACTAAAACCCAGATGGAACCAAGTCACTGCTCCCAAAAC	2736
2461	3b	CGGCATCAATGAATGGCTAACTAAAACCCAGATGGAACCAAGTCACTGCTCCCAAAAC	2520
2737	2y	TTCTGACTTGAGACCGAGAGACGTATCCAGACAGCGGAACTCTGGGAAAAGCAATCTGT	2796
2521	3b	TTCTGACTTGAGACCGAGAGACGTATCCAGACAGCGGAACTCTGGGAAAAGCAATCTGT	2580
2797	Qy	GGATAAGGTCACCTTCCCCACTTAAGGTTTGAGACAGTTCCAGAAAAGCCCAAGCTCAAG	2856
2581	3b	GGATAAGGTCACCTTCCCCACTTAAGGTTTGAGACAGTTCCAGAAAAGCCCAAGCTCAAG	2640
2857	Qy	ACGACGAGCAGCTCAGTTGTAGAGGCTAAATTCGCTCTGTTTTGTATTTATGTTGATTT	2916
2641	3b	ACGACGAGCAGCTCAGTTGTAGAGGCTAAATTCGCTCTGTTTTGTATTTATGTTGATTT	2700
2917	Qy	ACTAAATTTGGGTTCAATTATCTTTTAAATATCCAGTAAACCCATGATATATATC	2976
2701	3b	ACTAAATTTGGGTTCAATTATCTTTTAAATATCCAGTAAACCCATGATATATATC	2760
2977	2y	ACTATATTTAATTAATCAAGCTTAGAGATGTTCAATGTTAAAGTACTGCGCTTTGCAAGG	3036
2761	3b	ACTATATTTAATTAATCAAGCTTAGAGATGTTCAATGTTAAAGTACTGCGCTTTGCAAGG	2820
3037	2y	AGCTGTTTCTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCATAAAC	3096
2821	3b	AGCTGTTTCTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCATAAAC	2880
3097	Qy	TCGTGATCTGAGCAGTGATACCAACCAATCTGAAGTCAACAGAGATCCAAGTTTAAAA	3156
2881	3b	TCGTGATCTGAGCAGTGATACCAACCAATCTGAAGTCAACAGAGATCCAAGTTTAAAA	2940
3157	Qy	TTGCGTTCGCGAAATGTGTGAGTACTAGAAAAATGAACCGTAGTTTTTGTTTTTTAAAT	3216
2941	3b	TTG-CGCGAAATGTGTGAGTACTAGAAAAATGAACCGTAGTTTTTGTTTTTTAAAT	2999
3217	Qy	ACAGAGTCAATGTTTCTGCACTTTATAATAAGCATCGAGAAATATCTTAGTAGG	3276
3000	3b	ACAGAGTCAATGTTTCTGCACTTTATAATAAGCATCGAGAAATATCTTAGTAGG	3059
3277	Qy	CAATTTGAAACACTTTTTTGAAGTAGTAACCAATTTTCAAGTTTGAATCTGCGATATGTTG	3336

Db	3360	CAATGTGAACACTTTTGAAGTAACCCATTTCAGATTGAAATCTGCAATATGTTG	3119
Qy	3337	CTTTAAAAAAGAAAAAGATGCGGACAGTATTCTCTTT	3378
Db	3120	TCTTTAAAAAAGAAAAAGATGTAAGTAAAGTATTACTTTT	3161
RESULT 2			
AX336432			
LOCUS	AX336432	3610 bp	DNA
DEFINITION	Sequence 6941 from Patent WO0194629.		linear
ACCESSION	AX336432		
VERSION	AX336432.1	GI:18127151	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,		
JOURNAL	Horrigan, S., Soppet, D.R. and Weaver, Z.		
FEATURES	Cancer gene determination and therapeutic screening using signature		
source	gene sets		
ORIGIN	Patent: WO 0194629-A 6941 13-DEC-2001;		
	Avalon Pharmaceuticals (US)		
	Location/Qualifiers		
	1..3610		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
Query Match	84.2%;	Score 3116.4;	DB 6; Length 3610;
Best Local Similarity	99.3%;	Pred. No. 0;	
Matches 3140;	Conservative	0; Mismatches	21; Indels
			1; Gaps
Qy	217	CCTGCTGCGCCGCTTGTCTCTCTGCTGTGCTCTCTGCTTTAAAGAAATCAGTCTCTTC	276
Db	1	CCTGCTGCGCCGCTTGTCTCTCTGCTGTGCTCTCTGCTTTAAAGAAATCAGTCTCTTC	60
Qy	277	TTTCCGACTTAGTCTCTCGGGAAGAAGTTTCAGACTACAGAGTATCATTTGGAACATTTCAA	336
Db	61	TTTCCGACTTAGTCTCTCGGGAAGAAGTTTCAGACTACAGAGTATCATTTGGAACATTTCAA	120
Qy	337	GATCATCAAAATCCACAGGATTTGTGACCAACAGAGGCTCAGACATCTGATT	396
Db	121	GATCATCAAAATCCACAGGATTTGTGACCAACAGAGGCTCAGACATCTGATT	180
Qy	397	GCTGACCTGTCCAGACATCATCTGCTCTCCCTGAACTGAAATCACACCATGGATGATT	456
Db	181	GCTGACCTGTCCAGACATCATCTGCTCTCCCTGAACTGAAATCACACCATGGATGATT	240
Qy	457	TGAGCGTGCAGAGAACTTAGAGGCAAGAGGAGGAGATGCGACTCGAAGCAGAAAG	516
Db	241	TGAGCGTGCAGAGAACTTAGAGGCAAGAGGAGGAGATGCGACTCGAAGCAGAAAG	300
Qy	517	AATCCGCTACAGAGGAATGACGATGATGAAGAGGAGGACGCCCGGGAACCGCGCCGCG	576
Db	301	AATCCGCTACAGAGGAATGACGATGATGAAGAGGAGGACGCCCGGGAACCGCGCCGCG	360
Qy	577	AGCCGACAGGAAACCGCTCGGACAGAGGAGGAGGAATCTTTGGGACAGTGAACGA	636
Db	361	AGCCGACAGGAAACCGCTCGGACAGAGGAGGAGGAATCTTTGGGACAGTGAACGA	420
Qy	637	CCAGGTGAGGTGAATGCCAGAACAGTGTGCTGTGAGGAGGAGGCAAGAACACACAC	696
Db	421	CCAGGTGAGGTGAATGCCAGAACAGTGTGCTGTGAGGAGGAGGCAAGAACACACAC	480
Qy	697	AAACACTCAAGTGAAGGGGATGATGAGGCGGCAATCTCTGAGGCGCTGGCTCGGCGTGA	756
Db	481	AAACACTCAAGTGAAGGGGATGATGAGGCGGCAATCTCTGAGGCGCTGGCTCGGCGTGA	540

QY 757 GGAAGACGCCRAAAGCGCTTCAAGAGGCTCTGGAGCGGAGAGGAGTTGACCCCAAC 816
Db 541 GGAAGACGCCRAAAGCGCTTCAAGAGGCTCTGGAGCGGAGAGGAGTTGACCCCAAC 600
QY 817 AATAACAGATGCAAGTCTGCTGCCAAGCAGAGGAATGCAAAATGACACAGAGAAAA 876
Db 601 AATAACAGATGCAAGTCTGCTGCCAAGCAGAGGAATGCAAAATGACACAGAGAAAA 660
QY 877 TGAACCTACCGAGAGGAAGAAAAAGTGAAGTGCACCAAGAGATACGAGATAGAGA 936
Db 661 TGAACCTACCGAGAGGAAGAAAAAGTGAAGTGCACCAAGAGATACGAGATAGAGA 720
QY 937 AACAGAAAAAGTCAACCAAGTCTACCAAGAAATGATTCGAGGGATGCTGAAGAAAAA 996
Db 721 AACAGAAAAAGTCAACCAAGTCTACCAAGAAATGATTCGAGGGATGCTGAAGAAAAA 780
QY 997 GAAAGAGACAAAGAAAG 1056
Db 781 GAAAGAGACAAAGAAAG 840
QY 1057 AGAAATCAGGTAGAGTGTGTTGGAAGAGAAACAACTGAAAGCCAGGAGGAAACAGT 1116
Db 841 AGAAATCAGGTAGAGTGTGTTGGAAGAGAAACAACTGAAAGCCAGGAGGAAACAGT 900
QY 1117 GGTAAATGCTAAATAAATGGCGAGATCAGTTCAGAGAGGCTTAAACAAAGAGAGAGAG 1176
Db 901 GGTAAATGCTAAATAAATGGCGAGATCAGTTCAGAGAGGCTTAAACAAAGAGAGAGAG 960
QY 1177 GGAACAAGGTTTCAGATCAGATTTCCCATCATGAAAGATGGAAGAGGAGAGACAAAG 1236
Db 961 GGAACAAGGTTTCAGATCAGATTTCCCATCATGAAAGATGGAAGAGGAGAGACAAAG 1020
QY 1237 AGCTGAGGACAGAGGCAAGGTTGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1296
Db 1021 AGCTGAGGACAGAGGCAAGGTTGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1297 AGCAAAAAGATAGCAGATGAAACGACGAAGATTTGACGAGAGAGAGAGAGAGAGAG 1356
Db 1081 AGCAAAAAGATAGCAGATGAAACGACGAAGATTTGACGAGAGAGAGAGAGAGAGAG 1140
QY 1357 AGAAAG 1416
Db 1141 AGAAAG 1200
QY 1417 AGCAGAGGAGGACAGAGGATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1476
Db 1201 AGCAGAGGAGGACAGAGGATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1477 GATAAAGGAG 1536
Db 1261 GATAAAGGAG 1320
QY 1537 AAGGCGACAG 1596
Db 1321 AAGGCGACAG 1380
QY 1597 AGAGCAGAAACGTAACAGCAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1656
Db 1381 AGAGCAGAAACGTAACAGCAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1657 AAAAGGAG 1716
Db 1441 AAAAGGAG 1500
QY 1717 AGAAGATAAACTTCAGACAGCTGTCTTAAAGAAAACAGGAGAGAGAGAGAGAGAGAG 1776
Db 1501 AGAAGATAAACTTCAGACAGCTGTCTTAAAGAAAACAGGAGAGAGAGAGAGAGAGAG 1560
QY 1777 GCAAGCTTAAAG 1836
Db 1561 GCAAGCTTAAAG 1620
QY 1837 AGATGAAAAGATTAAAGAGGACAAAGAACCCCAAGAGAGAGAGAGAGAGAGAGAGAG 1896

Db 1621 AGATGAAAAGATTAAAGAGGACAAAGAACCCCAAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1897 AAAGAGAGGAGATTACAGAGTTAAGTCGAGAAATGGAGAAATTCATGACCCCAAACTTAA 1956
Db 1681 AAAGAGAGGAGATTACAGAGTTAAGTCGAGAAATGGAGAAATTCATGACCCCAAACTTAA 1740
QY 1957 ACATACGTAGAAATACTTTTCAGCCGCTTGGAGGAGGCGCCAGCGTGGACACCAAGAGGC 2016
Db 1741 ACATACGTAGAAATACTTTTCAGCCGCTTGGAGGAGGCGCCAGCGTGGACACCAAGAGGC 1800
QY 2017 TGAGGGCGCCCCCAGGTGGAAGCCGCAAAAGCTGGAGGAGTTCGTCGTCGCGCG 2076
Db 1801 TGAGGGCGCCCCCAGGTGGAAGCCGCAAAAGCTGGAGGAGTTCGTCGTCGCGCG 1860
QY 2077 GGAGACGAGAGCGAGAGAGTTCGAGAGAGCTCAACAGAGAGAGAGAGAGAGAGAGAGAG 2136
Db 1861 GGAGACGAGAGCGAGAGTTCGAGAGAGCTCAACAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 2137 GCTGGAGGAACTCAAGAAAAAGAGGAGAGAGAGAGAGAGAGTCTTGAGAGAGAGAGCA 2196
Db 1921 GCTGGAGGAACTCAAGAAAAAGAGGAGAGAGAGAGAGAGTCTTGAGAGAGAGAGCA 1980
QY 2197 GAG 2256
Db 1981 GAG 2040
QY 2257 AAAGGAGAGAGATTGAAAGGGGAGAGAGAGAGAGAGTCTGAGAGAGAGAGAGAGAG 2316
Db 2041 AAAGGAGAGATTGAAAGGGGAGAGAGAGAGAGAGTCTGAGAGAGAGAGAGAGAG 2100
QY 2317 AGATGGCTTTCAGATCAG 2376
Db 2101 AGATGGCTTTCAGATCAG 2160
QY 2377 CAAGATAG 2436
Db 2161 CAAGATAG 2220
QY 2437 ATGAGCCCATCAGCAGAGAGATAGTCTCCAGAGATTCAGAGAGAGAGAGAGAGAGAG 2496
Db 2221 ATGAGCCCATCAGCAGAGAGATAGTCTCCAGAGATTCAGAGAGAGAGAGAGAGAGAG 2280
QY 2497 TGCAATTTGAGGAG 2556
Db 2281 TGCAATTTGAGGAG 2340
QY 2557 TCCTGCTGAGAGGTTAGCGCAACATCAGAGATATGTTGGAGAGAGAGAGAGAGAGAGAG 2616
Db 2341 TCCTGCTGAGAGGTTAGCGCAACATCAGAGATATGTTGGAGAGAGAGAGAGAGAGAGAG 2400
QY 2617 CCCCACTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2676
Db 2401 CCCCACTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2460
QY 2677 CCGCATCAATGAATGGCTAACTAAAAACCCAGATGGAACAAAGTCACTGCTCCCAACC 2736
Db 2461 CCGCATCAATGAATGGCTAACTAAAAACCCAGATGGAACAAAGTCACTGCTCCCAACC 2520
QY 2737 TTCTGACTTGAGACACGAGAGAGAGATCCAGCAAGCGGAGAGAGAGAGAGAGAGAGAGAG 2796
Db 2521 TTCTGACTTGAGACACGAGAGAGAGATCCAGCAAGCGGAGAGAGAGAGAGAGAGAGAGAG 2580
QY 2797 GGATAAGGTCACTTCCCCCACTAAGGTTTGAGAGAGTTCAGAGAGAGAGAGAGAGAGAGAG 2856
Db 2581 GGATAAGGTCACTTCCCCCACTAAGGTTTGAGAGAGTTCAGAGAGAGAGAGAGAGAGAGAG 2640
QY 2857 ACGCAGAGCAGAGCTCAGTTGAGAGGCTAATTCGCTCTGTTTGTATTATGTTGATTT 2916
Db 2641 ACGCAGAGCAGAGCTCAGTTGAGAGGCTAATTCGCTCTGTTTGTATTATGTTGATTT 2700
QY 2917 ACTAAATTTGGGTTCAATATCTTTTATTTTCAATATCCAGTAAACCCATGTTATATTC 2976

QY	1537	AAGGCAGCAGAGGAGAGGCAAAAGGCGCAGGAGGAAGAAAGGCTAAGGTAGA	1596
Db	1321	AAGGCAGCAGAGGAGAGGCAAAAGGCGCAGGAGGAAGAAAGGCTAAGGTAGA	1380
QY	1597	AGAGCAGAAAACTGTAACAAGCAGCTAGAAGAGAAAAAATCATGCCATGCAAGAGACAAGAAT	1656
Db	1381	AGAGCAGAAAACTGTAACAAGCAGCTAGAAGAGAAAAAATCATGCCATGCAAGAGACAAGAAT	1440
QY	1657	AAAAGGGGAAAAGGTGAGACAGAAAAATAGAAGGGGAAATGGGTAAATGAAAAGAAAGCACA	1716
Db	1441	AAAAGGGGAAAAGGTGAGACAGAAAAATAGAAGGGGAAATGGGTAAATGAAAAGAAAGCACA	1500
QY	1717	AGAAGATAAACTTCACACAGCTCTCCTPAAAGAAAACAGGGAGAAAGGGAACCTAAAGT	1776
Db	1501	AGAAGATAAACTTCACACAGCTCTCCTPAAAGAAAACAGGGAGAAAGGGAACCTAAAGT	1560
QY	1777	GCAAGCTAAAAGAGAAAAAGCTCCAAAGAAAGCAAGCCTACCTTCAAAAAGAAAGAGATCAA	1836
Db	1561	GCAAGCTAAAAGAGAAAAAGCTCCAAAGAAAGCAAGCCTACCTTCAAAAAGAAAGAGATCAA	1620
QY	1837	AGATGAAAAGATTAAAAGGACAAAGAACCCAAAGAGAAAGTTAAGAGCTTTCATCGATCG	1896
Db	1621	AGATGAAAAGATTAAAAGGACAAAGAACCCAAAGAGAAAGTTAAGAGCTTTCATCGATCG	1680
QY	1897	AAAGAAAGGATTTCACAGAAAGTTAAGTCGCAGAAATGAGAAATTCATGACCCACAAACTTAA	1956
Db	1681	AAAGAAAGGATTTCACAGAAAGTTAAGTCGCAGAAATGAGAAATTCATGACCCACAAACTTAA	1740
QY	1957	ACATACTGAGATACHTTTCAGCGCGCTTCGAGGCGAGGCGCGTGGACACCAAGGAGGC	2016
Db	1741	ACATACTGAGATACHTTTCAGCGCGCTTCGAGGCGAGGCGCGTGGACACCAAGGAGGC	1800
QY	2017	TGAGGCGCCCCACGGTGAAAGCGCGCAAAAAGCTGGAGGAGCTTCGTGCTGCGCG	2076
Db	1801	TGAGGCGCCCCACGGTGAAAGCGCGCAAAAAGCTGGAGGAGCTTCGTGCTGCGCG	1860
QY	2077	GGAGACCGAGAGCGAAAGATTTCAGAAAGCTCAAAACAGAAAGCAGAGGAGCGGCTTTGGA	2136
Db	1861	GGAGACCGAGAGCGAAAGATTTCAGAAAGCTCAAAACAGAAAGCAGAGGAGCGGCTTTGGA	1920
QY	2137	GCTGAGGNACTTCAGAAAAGAGGGAGGAGAGAGAGAGTCTCTGGAGGAGGAGAGCA	2196
Db	1921	GCTGAGGNACTTCAGAAAAGAGGGAGGAGAGAGAGAGTCTCTGGAGGAGGAGAGCA	1980
QY	2197	GAGGAGGAACGAGGAGGAACCGATCGAAAACTCAGAGAGGAGGAGAGAGAGGCT	2256
Db	1981	GAGGAGGAACGAGGAGGAACCGATCGAAAACTCAGAGAGGAGGAGAGAGAGGCT	2040
QY	2257	RAAGGAAGAGATTGAAAGGCGMAAGACAGAGCTGCTCAGAAAAACCCAGAAAGATGCCAGA	2316
Db	2041	RAAGGAAGAGATTGAAAGGCGMAAGACAGAGCTGCTCAGAAAAACCCAGAAAGATGCCAGA	2100
QY	2317	AGATGCTGTGAGATGACAGAAAAACATTCAAGTGTGTTCACTCTCTAAAGTTTCATCTCT	2376
Db	2101	AGATGCTGTGAGATGACAGAAAAACATTCAAGTGTGTTCACTCTCTAAAGTTTCATCTCT	2160
QY	2377	CAAGATAGAAAGCGAGCAGAAATTTTTGAATTAAGTCTCTGCAGAAAAAGCAGTGGTGTCAA	2436
Db	2161	CAAGATAGAAAGCGAGCAGAAATTTTTGAATTAAGTCTCTGCAGAAAAAGCAGTGGTGTCAA	2220
QY	2437	ATCGACCCATCAGACGCAATAGTCTCCAAGATTGACAGACAGACTGGACAGTATATACCCAG	2496
Db	2221	ATCGACCCATCAGACGCAATAGTCTCCAAGATTGACAGACAGACTGGACAGTATATACCCAG	2280
QY	2497	TGCAATTGAGGGAAACAAAAAGCGCAAAACCTTACAAAGCGGCGAGCCTCGGATCTTCTGT	2556
Db	2281	TGCAATTGAGGGAAACAAAAAGCGCAAAACCTTACAAAGCGGCGAGCCTCGGATCTTCTGT	2340
QY	2557	TCTGCTGAGAGGTGATCGCAACATCAAGAGTATGTGGGAGAAAGGGAATGTGTTTTTCATC	2616
Db	2341	TCTGCTGAGAGGTGATCGCAACATCAAGAGTATGTGGGAGAAAGGGAATGTGTTTTTCATC	2400

QY	2617	CCCCATCGCAGCAGCAGCACCCAAATTAAGAAACTGCTGCCTTGAAGAGTAGGGTTTCTAG	2676
Db	2401	CCCCATCGCAGCAGCAGCAGCACCCAAATTAAGAAACTGCTGCCTTGAAGAGTAGGGTTTCTAG	2460
QY	2677	CGGCATCAATGAATGGCTAACTAAACCCCGAGATGGAACCAAGTCACTGCTCCCCAAACC	2736
Db	2461	CGGCATCAATGAATGGCTAACTAAACCCCGAGATGGAACCAAGTCACTGCTCCCCAAACC	2520
QY	2737	TTCTGACTTTGAGACACGAGAGAGCTATCCAGCAAGCGGAACCTCTCGGAAAGCAATCTGT	2796
Db	2521	TTCTGACTTTGAGACACGAGAGAGCTATCCAGCAAGCGGAACCTCTCGGAAAGCAATCTGT	2580
QY	2797	GGATAAGGTCACCTTCCCCCACTAAGTTTGAAGACAGTTCCAGAAAGAACCCAGCTCAAG	2856
Db	2581	GGATAAGGTCACCTTCCCCCACTAAGTTTGAAGACAGTTCCAGAAAGAACCCAGCTCAAG	2640
QY	2857	ACGCAGGACGAGCTCAGTTGTGAGAGGGCTAAATTCGGCTCTGTTTGTATTATCTTGATTT	2916
Db	2641	ACGCAGGACGAGCTCAGTTGTGAGAGGGCTAAATTCGGCTCTGTTTGTATTATCTTGATTT	2700
QY	2917	ACTAATAATGGGTTTCATTATCTTTTATTTTTCATATCCCGAGTAAACCCAGTATATTTATC	2976
Db	2701	ACTAATAATGGGTTTCATTATCTTTTATTTTTCATATCCCGAGTAAACCCAGTATATTTATC	2760
QY	2977	ACTATATTAATAAATCACAGCTCTAGAGATGTTTCATGGTAAAGTACTGCTTTCACAGG	3036
Db	2761	ACTATATTAATAAATCACAGCTCTAGAGATGTTTCATGGTAAAGTACTGCTTTCACAGG	2820
QY	3037	AGCCTGTTTCTAAAGAAACCCATCTGCTGTAAGTAAAGATCTTTTCTACTCATCATATAAC	3096
Db	2821	ATCCTGTTTCTAAAGAAACCCATCTGCTGTAAGTAAAGATCTTTTCTACTCATCATATAAC	2880
QY	3097	TCGTGATCTGAGCAGTGATACCAACACATCTGAAGTCAACAGAGATCCCAAGTTTAAAA	3156
Db	2881	TCGTGATCTGAGCAGTGATACCAACACATCTGAAGTCAACAGAGATCCCAAGTTTAAAA	2940
QY	3157	TTCCCTGCGGAATGTGTGCACTATCTAGAAAAATGAACCGTAGTTTGTGTTTTTAAAT	3216
Db	2941	TTG-CTGCGGAATGTGTGCACTATCTAGAAAAATGAACCGTAGTTTGTGTTTTTAAAT	2999
QY	3217	ACAGAAGTCAATGTTGTTTCTGCACTTTATATAAAGCATGCGAAGAAATATCTTAGTAGG	3276
Db	3000	ACAGAAGTCAATGTTGTTTCTGCACTTTATATAAAGCATGCGAAGAAATATCTTAGTAGG	3059
QY	3277	CAATTGTACACATTTTGAAGTAAACCCATTTTCAGATTTCGAATATCGCGATATGGTTG	3336
Db	3060	CAATTGTACACATTTTGAAGTAAACCCATTTTCAGATTTCGAATATCGCGATATGGTTG	3119
QY	3337	CTTTAAAAAAGAAAAAGATCGCGCAGTATTCCTTT	3378
Db	3120	TCCTTTAAAAAAGAAAAAGATGTACTGTTAGGTATTACTTTT	3161
RESULT 4			
HUMCD5M			
LOCUS	Human aorta caldesmon mRNA, complete cds.		
DEFINITION	Human aorta caldesmon mRNA, complete cds.		
ACCESSION	M83216		
VERSION	M83216.1 GI:306508		
KEYWORDS	caldesmon.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 3610)		
TITLE	Humphrey, M.B. Herrera-Sosa, H., Gonzalez, G., Lee, R. and Bryan, J.		
JOURNAL	Cloning of cDNAs encoding human caldesmons		
MEDLINE	Gene 112 (2), 197-204 (1992)		
PubMed	9220999		
COMMENT	On Jul 26, 1993 this sequence version replaced gi:180193.		
FEATURES	Original source text: Homo sapiens adult aorta cDNA to mRNA.		

```

1. .3610
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
/dev_stage="adult"
join(230..850,1616..2611)
/genes="caldesmon"
230..2611
/genes="caldesmon"
/genes="smooth muscle form"
/codon_start=1
/product="caldesmon"
/protein_id="AAA58419.1"
/db_xref="GI:180195"
/translation="MDPFRRLRQRKREMRLEARIAYQRNDDDEEEAARERRRR
ARQRRLRQKEESLQGVTDQVYNAQNSVDEEAKTTTNTQVGGDDDEAAFLERLAR
REERRQRKLQEARLQKEFQPTIDASLSLPSRRMQNDTAENETTEKEEKSRSORRY
EIEETETVTKSYQGNDRDAENKDEKEEKEEKEPKRGSI GENQVVMVEEKTTE
SQEETVSKNGQISNDEBEREQGSDIEISHKWEEDKAZAEARLEAEAE
REKAEODPKIADERAKIAEAEKAAQOEERREAEERREAEERKEKRAAEERQRIKEE
EYPAERQRIKEEKXAEAEERQRIKEEKZAAERQARAEAEERKAKVBEQKBNKOL
EEKRAERQRIKEEKVQKELEGWNEKKQAQEDKLQAVLKKQEESKGTQVQAKKEK
LQEDPTFFKKEEINKDEKIKDQKEEVEKVSFMDRKKGFTVKSQNGEFMHLKHTEN
TSRPGRASVDYTKAEAGQVQAGKRLLEELRRRKGTESEEFELKKQKQQAALLE
ELKKEERKRVLEEEQRKQAEADRLREEEKRRLKEEIERRAEAEAKRQKQME
DLGSDDKPKFCTPPKGSGLKIEBRAEFLNKSQKSSGVKSTHQAIAIVKSIDRLEQY
TSAIEGTSKAPKTPKAASDLPVPAEGVNRNISKMWKGNVFSSTAAGTENKGTAGLK
GVSRSINELWLTPTPDGKNKSPAPKPSDLRPGDVSSKRNLMWEKQSDVKTSPTKV"
join(230..850,1616..2611)
/genes="caldesmon"
/genes="non-muscle form"
/codon_start=1
/product="caldesmon"
/protein_id="AAA58420.1"
/db_xref="GI:180194"
/translation="MDQFRRLRQRKREMRLEARIAYQRNDDDEEEAARERRRR
ARQRRLRQKEESLQGVTDQVYNAQNSVDEEAKTTTNTQVGGDDDEAAFLERLAR
REERRQRKLQEARLQKEFQPTIDASLSLPSRRMQNDTAENETTEKEEKSRSORRY
EIEETETVTKSYQGNDRDAENKDEKEEKEEKEPKRGSI GENQIKDEKIKDQV
PKEEVKSFMDRKKGFTVKSQNGEFMHLKHTENFSRPGGRASVDYTKAEAGQVQE
AGKRLFEELRRRGTESEEFELKKQKQQAALLELKKRFERPKVLEERQQRKQIE
EADRLREEEKRRLKEIEIRRRRAEAEAKRQKMPEDGLSDDKKPKFCTPPKGSGLKIE
BRAEFLNKSQKSSGVKSTHQAIAIVKSIDRLEQYTSIAIEGTSKAPKTPKAASDLPV
PAEGVNRNISKMWKGNVFSSTAAGTENKGTAGLKRVGSSRSINELWLTPTPDGKNKSPAP
PSDLRPGDVSSKRNLMWEKQSDVKTSPTKV"
851..1615
intron
/genes="caldesmon"
/notes="non-muscle form"
3048
polya_site
3056
polya_site
3124
polya_site
3410
polya_site
3610
polya_site

```

ORIGIN

	Query Match	84.2%	Score 3116.4	DB 9	Length 3610	
	Best Local Similarity	99.3%	Pred. No. 0			
	Matches 3140	Conservative	0	Mismatches 21	Indels 1	Gaps 1
QY	217	CTGCTCTGCCGCGCTCTGCTCTCTGCGCTGTGCTCTGCTCTTAAAGAAATCATGCTCTTCC	276			
DB	1	CTGCTCTGCCGCGCTGCTCTCTGCTGTGCTCTGCTCTTAAAGAAATCATGCTCTTCC	60			
QY	277	TTTCCGACTTAGTCTCTCGGGAAGAAGTTTCAGACTACAAGGTATCATTTGAAACATTTCAA	336			
DB	61	TTTCCGACTTAGTCTCTCGGGAAGAAGTTTCAGACTACAAGGTATCATTTGAAACATTTCAA	120			

QY	337	GATCATCAAAATCAAAATTCACAGGGATTTGGTGAACCAACAGAGAGGCTCAGACATCTGATTT	396
DB	121	GATCA'CAAAATCAAAATTCACAGGGATTTGGTGAACCAACAGAGAGGCTCAGACATCTGATTT	180
QY	397	GCTGACCTGTCCAGACATCATCTGTGTCTCCCTGAACTTGAATTCACACCATGGATGATTT	456
DB	181	GCTGACCTGTCCAGACATCATCTGTGTCTCCCTGAACTTGAATTCACACCATGGATGATTT	240
QY	457	TGAGCGTCCAGAGAACTTAGAGGCAAAAGGGGAGGATCGACTCGAAGCAGAAAG	516
DB	241	TGAGCGTCCAGAGAACTTAGAAGGCAAAAGAGGAGGATCGACTCGAAGCAGAAAG	300
QY	517	AATCGCTTACAGAGAAATGACGATCATGAAGAGGAGCGCCGGGAACGGCGCCCGC	576
DB	301	AATCGCTTACAGAGAAATGACGATCATGAAGAGGAGCGCCGGGAACGGCGCCCGC	360
QY	577	AGCCCGACAGGAACGCTCGCGCAGAAAGCAGGAGGAAGATCTTGGGACAGGTGACCGA	636
DB	361	AGCCCGACAGGAACGCTCGCGCAGAAAGCAGGAGGAAGATCTTGGGACAGGTGACCGA	420
QY	637	CCAGTGTGAGGTGAATGCCCCAGAACTGAGTGTGCTGACAGGAGGCCCAAGACAACACAC	696
DB	421	CCAGTGTGAGGTGAATGCCCCAGAACTGAGTGTGCTGACAGGAGGCCCAAGACAACACAC	480
QY	697	AAACACTCAAGTGGAGGGATCATGAGGCCGATCTCTGAGCGCTGCTGCGCGTGA	756
DB	481	AAACACTCAAGTGGAGGGATCATGAGGCCGATCTCTGAGCGCTGCTGCGCGTGA	540
QY	757	GGAAAGCGCCAAAAACGCTTCAGAGGCTCTGAGCGGCGAGAGAGTTCGACCCAC	816
DB	541	GGAAAGCGCCAAAAACGCTTCAGAGGCTCTGAGCGGCGAGAGAGTTCGACCCAC	600
QY	817	AATPACAGATGCAAGTCTGTGCTCCCAAGCAGAGAAATGCAAAATGACACAGCAGAAAA	876
DB	601	AATPACAGATGCAAGTCTGTGCTCCCAAGCAGAGAAATGCAAAATGACACAGCAGAAAA	660
QY	877	TGAAACTACCGAGAGAGAAAAAGTGTGCGCAAGAGAAATGCGCAAGAAAGATACGAGATAGAGGA	936
DB	661	TGAAACTACCGAGAGAGAAAAAGTGTGCGCAAGAGAAATGCGCAAGAAAGATACGAGATAGAGGA	720
QY	937	AACAGAAAAACAGTCACCAAGTCTTACCAGAGAAATGATTTGGAGGGATGCTCTGAAAGAAAACAA	996
DB	721	AACAGAAAAACAGTCACCAAGTCTTACCAGAGAAATGATTTGGAGGGATGCTCTGAAAGAAAACAA	780
QY	997	GAAAGAGACAAGAAAAAGGAGAGAGAGAGAGAGAGCCAAAGCGGAGGAGCATTTGG	1056
DB	781	GAAAGAGACAAGAAAAAGGAGAGAGAGAGAGAGAGCCAAAGCGGAGGAGCATTTGG	840
QY	1057	AGAAAAACAGTACAGTGTGATGTTGCAAGAGAAAAACAATCTGAAAGCCAGAGAGAAAAACAGT	1116
DB	841	AGAAAAACAGTACAGTGTGATGTTGCAAGAGAAAAACAATCTGAAAGCCAGAGAGAAAAACAGT	900
QY	1117	GGTAATGTCTTAAAAATTTGGCGAGATCAGTTCAGAAAGAGCCCTTAAACAGAGAGGAGAG	1176
DB	901	GGTAATGTCTTAAAAATTTGGCGAGATCAGTTCAGAAAGAGCCCTTAAACAGAGAGGAGAG	960
QY	1177	GGAAACAAGTTCAGATGAGATTTCCCATCATGAAAAAGATGGAAGAGAGAAACAGAGAAAG	1236
DB	961	GGAAACAAGTTCAGATGAGATTTCCCATCATGAAAAAGATGGAAGAGAGAAACAGAGAAAG	1020
QY	1237	AGCTGAGGCGAGAGAGGCGAAGTTTGAACAGAGAAAGAGAAAGAAATTTAAAGCCGAGCA	1296
DB	1021	AGCTGAGGCGAGAGAGGCGAAGTTTGAACAGAGAAAGAGAAAGAAATTTAAAGCCGAGCA	1080
QY	1297	AGACAAAAAGATAGCAGATGAACCGACAGAAATTTGAAGCAGAGAAAAACAGCTGCCCA	1356
DB	1081	AGACAAAAAGATAGCAGATGAACCGACAGAAATTTGAAGCAGAGAAAAACAGCTGCCCA	1140
QY	1357	AGAAAGAAAGAGAGAGGCGAAGAGAGGGAAGAAAGATGAGGGAGGAGAAAGGGC	1416
DB	1141	AGAAAGAAAGAGAGAGGCGAAGAGAGGGAAGAAAGATGAGGGAGGAGAAAGGGC	1200
QY	1417	AGCAGAGGAGAGGCGCAGAGGATAAAGGAGGAGAGAAAAAGGCGCAGCAGAGAGGCGAGAG	1476

1141	Db	AGAAAGGAAGGAGAGAGGCGAAGAGAGAGGGAAGGATGAGGAGGAAGAGAAAAAGGC	1200
1417	Qy	AGCAGAGGAGGCGAGAGGATAAGGAGGAGAGAAAAAGGCGACGACGAGAGAGGCGAG	1476
1201	Db	AGCAGAGGAGGCGAGAGGATAAGGAGGAGAGAAAAAGGCGACGACGAGAGAGGCGAG	1260
1477	Qy	GATAAGGAGGAAGAGAAAAAGGCGACGAGAGGAGAGGCGACAGAGATAAAGAGAGAAAGAA	1536
1261	Db	GATAAGGAGGAAGAGAAAAAGGCGACGAGAGGAGAGGCGACAGAGATAAAGAGAGAAAGAA	1320
1537	Qy	AAGGGCAGCAGAGAGAGGCAAAAGGGCCAGGCGCAGAGAGAGAGAAAGGCTTAAGGTAGA	1596
1321	Db	AAGGGCAGCAGAGAGAGGCAAAAGGGCCAGGCGCAGAGAGAGAGAAAGGCTTAAGGTAGA	1380
1597	Qy	AGAGCGAARACGTAAACAGCAGCTAGAGAGAGAAARACATGCCATGCAAGAGCAAAAGT	1656
1381	Db	AGAGCGAARACGTAAACAGCAGCTAGAGAGAGAAARACATGCCATGCAAGAGCAAAAGT	1440
1657	Qy	AAAAGGGGAAAAAGGTAGAAACAGAAAAATAGAAAGGGAATGGGTAAATGAAAAGAAAGCAC	1716
1441	Db	AAAAGGGGAAAAAGGTAGAAACAGAAAAATAGAAAGGGAATGGGTAAATGAAAAGAAAGCAC	1500
1717	Qy	AGAAGATAAACTTCAGACAGCTGCTCTAAAGAAACAGGAGAGAGAGAGGGAACCTAAAGT	1776
1501	Db	AGAAGATAAACTTCAGACAGCTGCTCTAAAGAAACAGGAGAGAGAGAGGGAACCTAAAGT	1560
1777	Qy	GCAAGCTAAAGAGAAAAAGTCTCAAGAAAGCAAGCCTACCTTCAAAAAGAAAGAGATCAA	1836
1561	Db	GCAAGCTAAAGAGAAAAAGTCTCAAGAAAGCAAGCCTACCTTCAAAAAGAAAGAGATCAA	1620
1837	Qy	AGATGAAAAGATTAAAAGGACAAAGAACCCAAAGAAAGATTTAAGAGCTTCATGATGTCG	1896
1621	Db	AGATGAAAAGATTAAAAGGACAAAGAACCCAAAGAAAGATTTAAGAGCTTCATGATGTCG	1680
1897	Qy	AAAGAGGGATTTACAGAAGTAAAGTCGAGAAATGGAGAAATTCTATGACCACCAAACTTAA	1956
1681	Db	AAAGAGGGATTTACAGAAGTAAAGTCGAGAAATGGAGAAATTCTATGACCACCAAACTTAA	1740
1957	Qy	ACATACTGAGATAACTTTTACGCCCCCTTGAGGGAGGGCCAGCGTGACACCAAGGAGGC	2016
1741	Db	ACATACTGAGATAACTTTTACGCCCCCTTGAGGGAGGGCCAGCGTGACACCAAGGAGGC	1800
2017	Qy	TGAGGGCGCCCCCAGGTGGAAGCCGGCAAAAGGCTGGAGAGCTTCGTGCTGTCGCGG	2076
1801	Db	TGAGGGCGCCCCCAGGTGGAAGCCGGCAAAAGGCTGGAGAGCTTCGTGCTGTCGCGG	1860
2077	Qy	GGAGCCGAGACGGAAGAGTTTCGAGAAAGCTCAAAACAGACAGCAGAGAGCGGCTTTGGA	2136
1861	Db	GGAGCCGAGAGCGAAGATTTCGAGAAAGCTCAAAACAGACAGCAGAGAGCGGCTTTGGA	1920
2137	Qy	GCTGGAGGAATCAAGAAAAAAGAGGGAGGAGAGAAAGGTCTCTGAGAGAGGAAGCA	2196
1921	Db	GCTGGAGGAATCAAGAAAAAAGAGGGAGGAGAGAAAGGTCTCTGAGAGAGGAAGCA	1980
2197	Qy	GAGGAGGAAGCAGAGGGAAGCCGATCGAAAACCTCAGAGAGAGGAAACAGAAAGAGAGGCT	2256
1981	Db	GAGGAGGAAGCAGAGGGAAGCCGATCGAAAACCTCAGAGAGAGGAAACAGAAAGAGAGGCT	2040
2257	Qy	AAAGGAAGAGNTTGAAGGGCGAAGACAGAAAGCTGCTGAGAAAGCCAGAAAGATGCCAGA	2316
2041	Db	AAAGGAAGAGNTTGAAGGGCGAAGACAGAAAGCTGCTGAGAAAGCCAGAAAGATGCCAGA	2100
2317	Qy	AGATGGCTTGTTCAGATGACAGAAACCATTCAGAGTGTTCACCTCTAAAGGTTTCATCTCT	2376
2101	Db	AGATGGCTTGTTCAGATGACAGAAACCATTCAGAGTGTTCACCTCTAAAGGTTTCATCTCT	2160
2377	Qy	CAAGATAGAGAGCGAGCAGAAATTTTGAATTAAGTCTGTGCAGAAAAAGCATGGTGTCAA	2436
2161	Db	CAAGATAGAGAGCGAGCAGAAATTTTGAATTAAGTCTGTGCAGAAAAAGCATGGTGTCAA	2220
2437	Qy	ATCCAGCCCATCAAGCAGCAATAGTCTCTCCAGATTGACAGCAGACTGAGCAGATATACCAG	2496

Db	2221	ATCGA	CCCACTCAAGCAGCAATAAGTCTCCAGA	TTGACAGCAGACTGGAGCAGTATACCAAG	2289
Qy	2497	TGCAAT	TGAGGGAAACAAAAGCGCAAAACCTACAAG	CCGCGACCTCGGATCTTCTCTGT	2556
Db	2281	TGCAAT	TGAGGGAAACAAAAGCGCAAAACCTACAAG	CCGCGACCTCGGATCTTCTCTGT	2340
Qy	2557	TCTGCT	GTAAGGTGTACGCAACATCAAGAGTATGTGGG	AGAAAGGAATCTGTGTTTCATC	2616
Db	2341	TCTGCT	GTAAGGTGTACGCAACATCAAGAGTATGTGGG	AGAAAGGAATCTGTGTTTCATC	2400
Qy	2617	CCCACT	CTGCAGCAGGCACACCAAAATTAAGGAAATCTGCTGG	CTTGAAAGTATAGGGGTTTCTAG	2676
Db	2401	CCCACT	CTGCAGCAGGCACACCAAAATTAAGGAAATCTGCTGG	CTTGAAAGTATAGGGGTTTCTAG	2460
Qy	2677	CCGCAT	CAATGAATGGCTAACTAAAAACCCAGAGATGGA	AAACAAAGTCACTGCTCCCAAACC	2736
Db	2461	CCGCAT	CAATGAATGGCTAACTAAAAACCCAGAGATGGA	AAACAAAGTCACTGCTCCCAAACC	2520
Qy	2737	TTCTGACT	TGAGACCGAGAGACGTATCCAGCAACGGAACTCTGGG	AAAGCAATCTGT	2796
Db	2521	TTCTGACT	TGAGACCGAGAGACGTATCCAGCAACGGAACTCTGGG	AAAGCAATCTGT	2580
Qy	2797	GGATAAGGT	CACATTCGCCCACTAAGGTTTGAGACAGTTC	CAGAAAGCAAGCTCAAG	2856
Db	2581	GGATAAGGT	CACATTCGCCCACTAAGGTTTGAGACAGTTC	CAGAAAGCAAGCTCAAG	2640
Qy	2857	ACGACGACGAGCT	CAGTCTGTAGAGGCTAAATTCGCTCTGTTTG	TATTTATGTTGATTT	2916
Db	2641	ACGACGACGAGCT	CAGTCTGTAGAGGCTAAATTCGCTCTGTTTG	TATTTATGTTGATTT	2700
Qy	2917	ACTAAAT	TGCGGTTCAATATCTCTTTATTTTTC	CAATAATCCCAAGTAAACCCATGTATATTATC	2976
Db	2701	ACTAAAT	TGCGGTTCAATATCTCTTTATTTTTC	CAATAATCCCAAGTAAACCCATGTATATTATC	2760
Qy	2977	ACTATATTT	TAATAATCACAGCTCTAGAGATGTTTCATG	TGTAAAGTACTGCGCTTTGCACAGG	3036
Db	2761	ACTATATTT	TAATAATCACAGCTCTAGAGATGTTTCATG	TGTAAAGTACTGCGCTTTGCACAGG	2820
Qy	3037	AGCGTGT	TTCTAAAGAACCCATCTGTGTGAAATAGAGAC	TTTTTCTACTGATCATCATAAAC	3096
Db	2821	ATCCTGT	TTCTAAAGAACCCATCTGTGTGAAATAGAGAC	TTTTTCTACTGATCATCATAAAC	2880
Qy	3097	TCGTGAT	CTCAGCAGTGATACCAACACACATCTG	GAAGTCAACAGAAAGTCAAAAGTTAAAAA	3156
Db	2881	TCGTGAT	CTCAGCAGTGATACCAACACACATCTG	GAAGTCAACAGAAAGTCAAAAGTTAAAAA	2940
Qy	3157	TTGCGCT	CGGAAATGTGTGCAGTATCTAGAAAAATGA	ACCGTAGTTTTTGTGTTTTTTTAAAT	3216
Db	2941	TTG-CTG	CGGAAATGTGTGCAGTATCTAGAAAAATGA	ACCGTAGTTTTTGTGTTTTTTTAAAT	2999
Qy	3217	ACAGAAGT	CAATGTTTCTTGCACTTTTATATAAAGCATG	CGAAGAAATTTATCTTAGTAGG	3276
Db	3000	ACAGAAGT	CAATGTTTCTTGCACTTTTATATAAAGCATG	CGAAGAAATTTATCTTAGTAGG	3059
Qy	3277	CAAT	TGTAAACATTTTGAAGTAAACCCATTT	CAGATTTTCAAAATACTGCGATATGTTG	3336
Db	3060	CAAT	TGTAAACATTTTGAAGTAAACCCATTT	CAGATTTTCAAAATACTGCGATATGTTG	3119
Qy	3337	CTTTAAAA	AAAAAAGAAATCGCGCAGCAGTATT	TCCCTTT	3378
Db	3120	TC	TTTTAAAAAAGAAATGTACTGTTA	AGGTAATTACTTTTT	3161

RESULT 6						
HSM808959						
LOCUS	HSM808959	4091 bp	mRNA	linear	PRI 30-AUG-2003	
DEFINITION	Homo sapiens mRNA; CDNA DKFp686A06190 (from clone DKFp686A06190) .					
ACCESSION	BX648808					
VERSION	BX648808.1	GI:34367973				
KEYWORDS	.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					


```

>> 1482 ACTGACGAGCAGCACCACAAATAAGAAACTGCTGGCTTGAAGTAGGGGTTTCTAGCCGT 1541
>> 2681 ATCAATGAATGGCTAACTAAACCCAGATGGAAACAAAGTCACCTGCTCCCAACCTTCT 2740
>> 1542 ATCAATGAATGGCTAACTAAACCCAGATGGAAACAAAGTCACCTGCTCCCAACCTTCT 1601
>> 2741 GACTTGAGACAGAGAGAGCTATCCAGCAGCGGAAACCTCTGGGAAAGCAATCTGTGGAT 2800
>> 1602 GACTTGAGACAGAGAGAGCTATCCAGCAGCGGAAACCTCTGGGAAAGCAATCTGTGGAT 1661
>> 2801 AAGTCTACTTCCCACTAAGTCTTGGACAGTTCAGAAAGCAACCAAGCTCAAGACGC 2860
>> 1662 AAGTCTACTTCCCACTAAGTCTTGGACAGTTCAGAAAGCAACCAAGCTCAAGACGC 1721
>> 2861 AGGACGAGCTCAGTGTAGAGGGCTAAATCGCTGCTGTTTGTATTTATTTGTTTACTA 2920
>> 1722 AGGACGAGCTCAGTGTAGAGGGCTAAATCGCTGCTGTTTGTATTTATTTGTTTACTA 1779
>> 2921 AATTGGGTTCAATATCTTTTATTTTCAATATCCCAAGTAAACCCATGTATATTACTA 2980
>> 1780 AATTGGGTTCAATATCTTTTATTTTCAATATCCCAAGTAAACCCATGTATATTACTA 1839
>> 2981 TATTTAATATCAGAGCTAGAGATGTTTCATGTTAAAGTACTGCTTTTGCACAGGAGCC 3040
>> 1840 TATTTAATATCAGAG-CTAGAGATGTTTCATGTTAAAGTACTGCTTTTGCACAGGAGCC 1898
>> 3041 TGTCTTAAAGAAACCCATGCTGTGAAATPAGAGACTTTTCTACTGATCATATAAATCTTG 3100
>> 1899 TGTCTTAAAGAAACCCATGCTGTGAAATPAGAGACTTTTCTACTGATCATATAAATCTTG 1958
>> 3101 TATCTGAGCAGTATACCAACACATCTGAAGTCAACAGAGATCCAAAGTTTAAATATGC 3160
>> 1959 TATCTGAGCAGTATACCAACACATCTGAAGTCAACAGAGATCCAAAGTTTAAATATGC 2018
>> 3161 CTGCGGAATGTGTGAGTATCTAGAAAATGAACCGTAGTTTGTGTTTAAATATACAG 3220
>> 2019 CTGCGGAATGTGTGAGTATCTAGAAAATGAACCGTAG-TTTTGTGTTTAAATATACAG 2077
>> 3221 AAGTCATGTTGTTCTGCACTTTATAATAAGCATGGAAGAAATTTATCTTAGTAGCAAT 3280
>> 2078 AAGTCATGTTGTTCTGCACTTTATAATAAGCATGGAAGAAATTTATCTTAGTAGCAAT 2137
>> 3281 TGTAAACACTTTTGAAGTAAACCCATTTGAAATTTGAAATCTGCGATATGTTGTTT 3340
>> 2138 TGTAAACACTTTTGAAGTAAACCCATTTGAAATTTGAAATCTGCAATATGTTGTTT 2197
>> 3341 AAAAAAAAAAAGAAAGA 3357
>> 2198 TAAAAAAAAAAAAAAAAA 2214

RESULT 8
LOCUS E05382 2215 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA encoding human caldesmon.
ACCESSION E05382
VERSION E05382.1 GI:2173571
KEYWORDS JP 1993230099-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
HAYASHI, K., Hashida, H., Asada, K., Kotani, H., Katou, I. and Sofue, K.
POLYPEPTIDE
PATENT: JP 1993230099-A 3 07-SEP-1993;
SOFUE KENJI, TAKARA SHUZO CO LTD
OS Homo sapiens (human)
PN JP 1993230099-A/3
PD 07-SEP-1993
PF 27-DEC-1991 JP 1991358040
PR 29-MAR-1991 JP 91P 89106
PI HAYASHI KENICHIRO, HASHIDA HISASHI, ASADA KIYOZOU, PI KOTANI

```

```

HIROICHI,
PI KATOU IKUNOSHIN, SOFUE KENJI
PC C07K13/00,C12N15/12//A61K37/02,C12P21/02,G01N33/53,(C12P21/02,
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC *source: cell_type=Hela;
FH Key Location/Qualifiers
FH CDS 12..1688
FT polyA signal /product='human caldesmon'
FT 2103..2108
FT polyA site 2199..2215.
FEATURES
source 1..2215
/organism='Homo sapiens'
/mol_type='genomic RNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 42.4%; Score 1570.2; DB 6; Length 2215;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 93; Indels 4; Gaps 3;

Qy 1601 CAGAAACCTAACAGCAGCTAGAGAGAAACATGCTCCATGCAAGAGACAAAGATAAAA 1660
Db 462 CAAGAAAGATACGAGATAGAGGAAACAGAAACAGTCACCAAGTCCTACCAAGAGATGAT 521
Qy 1661 GGGGAAAGGTAGAACAGAAATAGAGGGAAATGGTAAATGAAAAAGAAAGCAAGAA 1720
Db 522 TGGAGGGATGCTGAGAGAAACCAAGAAAGAAAGCAAGAAAGGAGGAGAGAGAG 581
Qy 1721 GATAAATTCAGACAGCTGCTCTAAAGAAACAGGGAGAGAGAGAGAACTTAAAGTGCAA 1780
Db 582 AAGCAAAAGCGAGGAGCATTTGGAGAAATTCAGGGAGAGAGAGAGAACTTAAAGTGCAA 641
Qy 1781 GCTAAACAGAGAAAGCTCCAGAGACAGCCCTTACCTTCAAAAAAGAGAGATCAAAAGAT 1840
Db 642 GCTAAACAGAGAAAGCTCCAGAGAGACAGCCCTTACCTTCAAAAAAGAGAGATCAAAAGAT 701
Qy 1841 GAAAGAGTTAAAAAGCAGAAAGAACCCAAAGAAAGTTAAGAGCTTTCATGATCGAAAG 1900
Db 702 GAAAGAGTTAAAAAGCAGAAAGAACCCAAAGAAAGTTAAGAGCTTTCATGATCGAAAG 761
Qy 1901 AAGGATTTACAGAGTTAAGTCGAGAGTGGAGATTCATGACCCCAACATTTAAACAT 1960
Db 762 AAGGATTTACAGAGTTAAGTCGAGAGTGGAGATTCATGACCCCAACATTTAAACAT 821
Qy 1961 ACTGAGAAATCTTTTCAGCGCCCTTCGAGGAGGCGCCAGCTGGACACCAAGGAGCTGAG 2020
Db 822 ACTGAGAAATCTTTTCAGCGCCCTTCGAGGAGGCGCCAGCTGGACACCAAGGAGCTGAG 881
Qy 2021 GCGCCCCCAGGTGGAAGCCGCGCAAAAGGCTGGAGAGGTTCTGCTCGTCGCGGGAG 2080
Db 882 GCGCCCCCAGGTGGAAGCCGCGCAAAAGGCTGGAGAGGTTCTGCTCGTCGCGGGAG 941
Qy 2081 ACCGAGAGCGAAGAGCTTCGAGAAGCTCAAACAGAGCAGCAGGAGCGGCTTTGGAGCTG 2140
Db 942 ACCGAGAGCGAAGAGCTTCGAGAAGCTCAAACAGAGCAGCAGGAGCGGCTTTGGAGCTG 1001
Qy 2141 GAGGACTCAGAAAAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2200
Db 1002 GAGGAACTCAAAGAAAAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1061
Qy 2201 AGGAAGCAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2260
Db 1062 AGGAAGCAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1121
Qy 2261 GAAGAGATTGAAGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2320
Db 1122 GAAGAGATTGAAGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1181
Qy 2321 GGCTGTTCAGATGACAAAGAAACCAATTCAGAGTGTTCCTCTCTAAAGGTTTCATCTCTCAAG 2380

```

1152 GCGTTGTGATGACAGAAACCACTTCAAGTGTTCATCTCTAAAGTTCATCTCTCAAG 1241
 2381 ATAGAAGAGGACAGCAATTTTGAATAGTCTGTGAGAAAGCAGTGGTGTCAATCG 2440
 1242 ATAGAAGAGGACAGCAATTTTGAATAGTCTGTGAGAAAGCAGTGGTGTCAATCG 1301
 2441 ACCATCAAGAGCAATAGTCTCAAGATTGACAGAGAGCTGGAGCAGTATACAGTGCA 2500
 1302 ACCATCAAGAGCAATAGTCTCAAGATTGACAGAGAGCTGGAGCAGTATACAGTGCA 1361
 2501 ATTGAGGGAACAAAAGCGCAAACTTACAAAGCGCAGCTCGATCTTCTCTCTCT 2560
 1362 ATTGAGGGAACAAAAGCGCAAACTTACAAAGCGCAGCTCGATCTTCTCTCTCT 1421
 2561 GCTGAAGTGTACGCAATCAAGAGTATGTGGAGAAAGGGAATGTGTTTTCATCTCC 2620
 1422 GCTGAAGTGTACGCAATCAAGAGTATGTGGAGAAAGGGAATGTGTTTTCATCTCC 1481
 2621 ACTGACAGGACACCAATATAGGAACCTGCTGGCTTGAAGGTAGGGTTCCTAGCCG 2680
 1482 ACTGACAGGACACCAATATAGGAACCTGCTGGCTTGAAGGTAGGGTTCCTAGCCG 1541
 2681 ATCAATGAATGGCTAACTAAACCCAGATGGAACCAAGTCACTGCTCCCAACCTTCT 2740
 1542 ATCAATGAATGGCTAACTAAACCCAGATGGAACCAAGTCACTGCTCCCAACCTTCT 1601
 2741 GACTTGACAGGACGATCCAGAGGGAACCTCTGGGAAGCAATCTGTGGAT 2800
 1602 GACTTGACAGGACGATCCAGAGGGAACCTCTGGGAAGCAATCTGTGGAT 1661
 2801 AAGGTCACCTTCCCACTAAGGTTTGAGCAGTTCAGAAAGCAACCAAGCTCAAGACG 2860
 1662 AAGGTCACCTTCCCACTAAGGTTTGAGCAGTTCAGAAAGCAACCAAGCTCAAGACG 1721
 2861 AGGACAGCTCAGTTGAGAGGCTATTCGCTGTTGTTATTTATGTTGATTTACTA 2920
 1722 AGGACAGCTCAGTTGAGAGGCTATTCGCTGTTGTTATTTATGTTGATTTACTA 1779
 2921 AATTGGGTTCAATTATCTTTTATTTTCAATATCCAGTAAACCCATGATATATCACTA 2980
 1780 AATTGGGTTCAATTATCTTTTATTTTCAATATCCAGTAAACCCATGATATATCACTA 1839
 2981 TATTTAATATCAAGTCTAGAGATGTTCTGTTAAAGTACTGCTTTGACAGAGGCC 3040
 1840 TATTTAATATCAAG-CTAGAGATGTTCTGTTAAAGTACTGCTTTGACAGAGGCC 1898
 3041 TGTCTTAAAGAAACCCATGCTGTGAATAGAGACTTTTCTACTGATCATCACTACTG 3100
 1899 TGTCTTAAAGAAACCCATGCTGTGAATAGAGACTTTTCTACTGATCATCACTACTG 1958
 3101 TATCTGAGCAGTATACCAACCCATGCTGAAGTCAACAGAGAGTCCAAAGTTTAAATTCG 3160
 1959 TATCTGAGCAGTATACCAACCCATGCTGAAGTCAACAGAGAGTCCAAAGTTTAAATTCG 2018
 3161 CTGCGAAATGTTGAGTATCTAGAAATTAAGACCGTAGTTTGTGTTTAAATACAG 3220
 2019 CTGCGAAATGTTGAGTATCTAGAAATTAAGACCGTAG-TTTGTTTAAATACAG 2077
 3221 AAGTCATGTTTCTGCTTATATAAAGCATGGAAGAAATATCTTAGTAGGCAAT 3280
 2078 AAGTCATGTTTCTGCTTATATAAAGCATGGAAGAAATATCTTAGTAGGCAAT 2137
 3281 TGTAAACACTTTTGAAGTAACCATTTGAGTTTGAATTAAGTAAATGTTGTTT 3340
 2138 TGTAAACACTTTTGAAGTAACCATTTGAGTTTGAATTAAGTAAATGTTGTTT 2197
 3341 AAAAAAAGAAAAA 3357
 2198 TAAAAAAGAAAAA 2214

RESULT 9

123421
 LOCUS 2215 bp DNA linear PAT 07-OCT-1996
 DEFINITION Sequence 14 from patent US 5532337.
 ACCESSION 123421
 VERSION 123421.1 GI:1603291
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 UNCLASSIFIED.
 REFERENCE 1 (bases 1 to 2215)
 AUTHORS Hayashi,K., Hashida,T., Asada,K., Kotani,H., Kato,I. and Sobue,K.
 TITLE Proteins comprising calmodulin- and actin-binding human caldesmon
 JOURNAL Polypeptide fragments
 FEATURES Patent: US 5532337-A 14 02-JUL-1996;
 Location/Qualifiers
 1. .2215
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Query Match 42.4%; Score 1570.2; DB 6; Length 2215;
 Best Local Similarity 94.5%; Pred. No. 0;
 Matches 1660; Conservative 0; Mismatches 93; Indels 4; Gaps 3;
 QY 1601 CAGAAACGTAACAGCAGCTAGAGAGAAACATGCCATGCAAGAGACAAAGATAAAA 1660
 DB 462 CAAGAAAGATACGATAGAGGAAACAGAAACAGTCCCAAGTCTTACCCAGAAATGAT 521
 QY 1661 GGGAAAGGTAGACAGAAATAGAGGGAATGGTAAATGAAGAAAGACACAGAA 1720
 DB 522 TGGAGGATGCTGAAGAAACAAAGAAAGACAAAGGAAAGGAGGAGGAGAGAG 581
 QY 1721 GATAAACTTCAGACAGCTGCTTAAAGAAACAGGGAGAGAGAGGAACTTAAAGTGCAA 1780
 DB 582 AGCCAAAGCGGGAGCAATTGGAGAAATCAGGGAGAGAGAGGAACTTAAAGTGCAA 641
 QY 1781 GCTAAAGAGAAAGCTCCAGAAAGACAGCTACCTTCAAAAAGAGAGAGATCAAAGAT 1840
 DB 642 GCTAAAGAGAAAGCTCCAGAAAGACAGCTACCTTCAAAAAGAGAGATCAAAGAT 701
 QY 1841 GAAAGATTAAAAAGGACAAAGAACCCAAAGAAAGTAAAGAGCTTCATGGATCGAAAG 1900
 DB 702 GAAAGATTAAAAAGGACAAAGAACCCAAAGAAAGTAAAGAGCTTCATGGATCGAAAG 761
 QY 1901 AAGGATTTACAGAGTTAAGTGCAGAAATGAGAAATTCATGACCCACAACTTAAACAT 1960
 DB 762 AAGGATTTACAGAGTTAAGTGCAGAAATGAGAAATTCATGACCCACAACTTAAACAT 821
 QY 1961 ACTGAGAACTTTTACGCGCCCTTGGAGGGAGGCGCAGCTGGACACCAAGAGGCTGAG 2020
 DB 822 ACTGAGAACTTTTACGCGCCCTTGGAGGGAGGCGCAGCTGGACACCAAGAGGCTGAG 881
 QY 2021 GCGCCCCCAGGTGGAAGCGGCAAAAGGCTGGAGGAGCTTCGTCGTCGTCGCGGGAG 2080
 DB 882 GCGCCCCCAGGTGGAAGCGGCAAAAGGCTGGAGGAGCTTCGTCGTCGTCGCGGGAG 941
 QY 2081 ACCGAGAGGAGAGTTTCGAGAAAGTCAACAGAGAGCAGGAGGCGGCTTTGAGAGCTG 2140
 DB 942 ACCGAGAGGAGAGTTTCGAGAAAGTCAACAGAGCAGCAGGAGGCGGCTTTGAGAGCTG 1001
 QY 2141 GAGGAACTCAAGAAAAGAGGGAGAGAGAGAAAGGAGTCTTGGAGGAGAGAGAGAGG 2200
 DB 1002 GAGGAACTCAAGAAAAGAGGGAGAGAGAGAAAGGAGTCTTGGAGGAGAGAGAGAGG 1061
 QY 2201 AGGAACGAG 2260
 DB 1062 AGGAACGAG 1121
 QY 2261 GAAGAGATTGAAAGGAT 2320
 DB 1122 GAAGAGATTGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1181
 QY 2321 GCGTTGTCAGATGACAGAAACCAATTCAAGTGTTCCTTCAAGAGGTTCTCTCTCAAG 2380

1182 GCCTTGTGAGTACAGAGAACCAATTCAGTGTTCCTCCTTAAGGTTCACTCTCAAG 1241
2381 ATAGAAGAGCAGCAGAGAAATTTTGAATAAGTCTGTGCAGAAAAGCAGTGTGTCAAAATCG 2440
1242 ATAGAAGAGCAGCAGAGAAATTTTGAATAAGTCTGTGCAGAAAAGCAGTGTGTCAAAATCG 1301
2441 ACCCATCAAGCAGCAGTGTCTCCAGATTGACAGCAGAGCTGGAGCAGTATACAGTGCA 2500
1302 ACCCATCAAGCAGCAGTGTCTCCAGATTGACAGCAGAGCTGGAGCAGTATACAGTGCA 1361
2501 ATTGAGGGAACAAAAGCGCAAAACCTACAAAGCGCGACCTCGATCTTCTGTGTTCT 2560
1362 ATTGAGGGAACAAAAGCGCAAAACCTACAAAGCGCGACCTCGATCTTCTGTGTTCT 1421
2561 GCTGAGGTGTACGCAACATCAAGAGTATGTGGAGAAAGGGAATGTGTTTATCCTCC 2620
1422 GCTGAGGTGTACGCAACATCAAGAGTATGTGGAGAAAGGGAATGTGTTTATCCTCC 1481
2621 ACTGCAGCAGCAGCACACCAATTAAGCAAACTGCTGGCTTGAAGTGAAGGTTTCTAGCCGC 2680
1482 ACTGCAGCAGCAGCACACCAATTAAGCAAACTGCTGGCTTGAAGTGAAGGTTTCTAGCCGT 1541
2681 ATCAATGATGCTTAATTAACCCAGATGGAACAGTCACTGCTCCCAACCTTCT 2740
1542 ATCAATGATGCTTAATTAACCCAGATGGAACAGTCACTGCTCCCAACCTTCT 1601
2741 GACTTGACACGAGGAGACGTATCCAGCAAGCGAAGCTCTGGGAAAGCAATCTGTGGAT 2800
1602 GACTTGACACGAGGAGACGTATCCAGCAAGCGAAGCTCTGGGAAAGCAATCTGTGGAT 1661
2801 AAGTCACTTCCCCCACTAAGTTTGAGACAGTTCCAGAAAGCAACCAAGCTCAAGACGC 2860
1662 AAGTCACTTCCCCCACTAAGTTTGAGACAGTTCCAGAAAGCAACCAAGCTCAAGACGC 1721
2861 AGGACGAGCTCAGTTGTAGAGGCTAATTCGCTCTGTTTGTATTTATGTTGTTTACTA 2920
1722 AGGACGAGCTCAGTTGTAGAGGCTAATTCGCTC--TTTGTATTTATGTTGTTTACTA 1779
2921 AATTGGGTTCAATATCTTTTATTTTCAATATCCAGTAAACCCATGTATATTATCACTA 2980
1780 AATTGGGTTCAATATCTTTTATTTTCAATATCCAGTAAACCCATGTATATTATCACTA 1839
2981 TATTTAATATCACAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTCACAGGAGCC 3040
1840 TATTTAATATCACAG-CTAGAGATGTTTCATGTTAAAGTACTGCTTTCACAGGAGCC 1898
3041 TGTCTTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCAATCACTCTG 3100
1899 TGTCTTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCAATCACTCTG 1958
3101 TATCTGAGCAGTGATACCAACACATCTGAAGTCAACAGAAATCCAAAGTTTAAATTC 3160
1959 TATCTGAGCAGTGATACCAACACATCTGAAGTCAACAGAAATCCAAAGTTTAAATTC 2018
3161 CTGCGGAATGTGTCAGTATCTAGAAAATGAACCGTAGTTTGTGTTTAAATACAG 3220
2019 CTGCGGAATGTGTCAGTATCTAGAAAATGAACCGTAG-TTTTGTGTTTAAATACAG 2077
3221 AAGTCATGTTGTTCTGCACTTTATTAATAAGCATGGAAGAAATTAATCTTAGTAGGCAAT 3280
2078 AAGTCATGTTGTTCTGCACTTTATTAATAAGCATGGAAGAAATTAATCTTAGTAGGCAAT 2137
3281 TGTAAACCTTTTGAAGTAACCCATTCAGATTGGAATGCAATGCTGATGTTGCTTT 3340
2138 TGTAAACCTTTTGAAGTAACCCATTCAGATTGGAATGCAATGCTGATGTTGCTTT 2197
3341 AAAAAAAAAAGAAAAAGA 3357
2198 TAAAAAAAAAAAAAAAAA 2214

RESULT 10

BC040354
LOCUS
DEFINITION Homo sapiens caldesmon 1, transcript variant 3, mRNA (cdna clone
ACCESSION BC040354
VERSION BC040354.1
KEYWORDS GI:25955565
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS
1 (bases 1 to 4130)
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Alschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, P.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scheetle, M., Soares, M.B., Bonaldo, M.F., Casavant, T.D.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fanech, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krawinski, M.I., Skalska, U., Smalish, D.E.,
Schnerch, A., Schain, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4130)
Strausberg, R.
Direct Submission
Submitted (27-NOV-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgpsb-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lousseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

FEATURES
source
1. 4130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:21352 IMAGE:4753285"
/tissue_type="Skin, squamous cell carcinoma"
/clone_lib="NCI CGAP_Skn4"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6.1"
1. 4130

gene

BC040354 4130 bp mRNA linear PRI 07-OCT-2003
Homo sapiens caldesmon 1, transcript variant 3, mRNA (cdna clone
MGC:21352 IMAGE:4753285), complete cds.
BC040354
BC040354.1 GI:25955565
MGC.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 4130)
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Alschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, P.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scheetle, M., Soares, M.B., Bonaldo, M.F., Casavant, T.D.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fanech, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krawinski, M.I., Skalska, U., Smalish, D.E.,
Schnerch, A., Schain, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4130)
Strausberg, R.
Direct Submission
Submitted (27-NOV-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgpsb-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lousseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAC Plate: 28 Row: p Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 15149468.
Location/Qualifiers
1. 4130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:21352 IMAGE:4753285"
/tissue_type="Skin, squamous cell carcinoma"
/clone_lib="NCI CGAP_Skn4"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6.1"
1. 4130

```

/genes="CALD1"
/notes="synonyms: CDM, NAG22"
/db_xref="LocusID:800"
/db_xref="MIM:114213"
202..1818
/codon_start=1
/product="caldesmon 1, isoform 3"
/protein_id="AAH40354.1"
/db_xref="GI:25955666"
/db_xref="LocusID:800"
/translations="MDFFERRELRRLRQKREMRLEARIYQRNDDBEEAARRRRR
ARQRRLQKQEBESLQVTOQVEVNAQNSVPEBAKTITNTQYEGDDAAFLERLAR
REERKRQLEALRQKEPDTTIDASLSPSRMONDTAENETTEKEKESRQERY
ETEEETVTKSYOKNRDRAENKEDKEKEEERKEKSGS IGENOIKDEKIKDKE
PKEEVKSPDRKKGFTVKSQNGEPMTKLKHIENTSRPGGRASVTKKEAGAPQVE
AKRULERRRGETSEEFELKQKQCEALSLLEELKKGREERKVLLEEEQRKQE
EADRKLREERERLKEBERRRAAEKQKMPEDGLSDDKKPKFCTPPKSSLUKE
ERAEFLNKSVQKSSGVKSTHQAIVSKIDSLRLEYTSAIEGTSKAKTPKPAASDLFVP
AEGVRNIGSMWEXGNVSSPTAAGTGNKETAAGLKVGVSRIINELWLTTPDGNKSPAPK
PSDLRPGDVSSKENLWEKQKQVDTKVTPTKV"
202..735
/misc_feature
/notes="Caldesmon"
/db_xref="CDD:pfam02029"
874..1038
/misc_feature
/notes="Caldesmon"
/db_xref="CDD:pfam02029"
1294..1815
/misc_feature
/notes="Caldesmon"
/db_xref="CDD:pfam02029"

ORIGIN
Query Match 42.2%; Score 1562; DB 9; Length 4130;
Best Local Similarity 75.2%; Pred. No. 0;
Matches 2340; Conservative 0; Mismatches 5; Indels 765; Gaps 1;
QY 245 CTGTGCTCTGCTTAAAGAAATCAGTCTCTTCCGACTTAGTCTCCGGGAAGAAGTT 304
DB 1 CTGTGCTCTGCTTAAAGAAATCAGTCTCTTCCGACTTAGTCTCCGGGAAGAAGTT 60
QY 305 TCAGACTACAGGTATCATTTGGAACATTTCAAGATCATCAATCAAAATCCACAGGGATT 364
DB 61 TCAGACTACAGGTATCATTTGGAACATTTCAAGATCATCAATCAAAATCCACAGGGATT 120
QY 365 GGTGACCAACCAAGGCTCAGACATCTGATGCTGACCTGCTCCAGACATCATCTGTCT 424
DB 121 GGTGACCAACCAAGGCTCAGACATCTGATGCTGACCTGCTCCAGACATCATCTGTCT 180
QY 425 CCTGAACTGAAATCACACCATGATGATTTTGGCGTCCAGAGAACTTAGAAGGCAA 484
DB 181 CCTGAACTGAAATCACACCATGATGATTTTGGCGTCCAGAGAACTTAGAAGGCAA 240
QY 485 AAGAGGGAGAGATGCGACTCGAAGCAGAAAGAAATCGCTTACCAGAGGAATGACGATGAT 544
DB 241 AAGAGGGAGAGATGCGACTCGAAGCAGAAAGAAATCGCTTACCAGAGGAATGACGATGAT 300
QY 545 GAAGAGAGGACGCGCGGGAACGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 604
DB 301 GAAGAGAGGACGCGCGGGAACGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 605 CAGGAGGAAGAAATCTTTGGGACAGGTGACCCAGCAGGTGGAGGTGAATGCCCAGAACAGT 664
DB 361 CAGGAGGAAGAAATCTTTGGGACAGGTGACCCAGCAGGTGGAGGTGAATGCCCAGAACAGT 420
QY 665 GTGCTTGACGAGAGGCGCAAGAACACCAACCAACCAACTCAAGTGGAGGGGATGATGAG 724
DB 421 GTGCTTGACGAGAGGCGCAAGAACACCAACCAACCAACTCAAGTGGAGGGGATGATGAG 480
QY 725 GCGGCATCTCTGAGCGCTTGGCTCGGCGTGGAGAAAGACGCCAAAAACGCCCTTCAGGAG 784
DB 481 GCGGCATCTCTGAGCGCTTGGCTCGGCGTGGAGAAAGACGCCAAAAACGCCCTTCAGGAG 540
QY 785 GCTCTGAGCGGGAAGAGGAGTTTCAGCCCAACAATTAACAGATGCAAGTCTGTCTGCTCCA 844
DB 844 GCTCTGAGCGGGAAGAGGAGTTTCAGCCCAACAATTAACAGATGCAAGTCTGTCTGCTCCA

```

```

DB 541 GCTCTGAGCGGCGGAGAGGAGTTGACCCCAACAATAACAGATGCAAGTCTGTCTGCCA 600
QY 845 AGCAGAAGAATGCAAAATGACACAGCAGAGAAAATGAACTACCGAGAGGAGGAGAAAAGT 904
DB 601 AGCAGAAGAATGCAAAATGACACAGCAGAGAAAATGAACTACCGAGAGGAGGAGAAAAGT 660
QY 905 GAAAGTCCGCAAGAAAGATACGAGATAGAGGAAACAGAAACAGTACCAAGTCTTACCAG 964
DB 661 GAAAGTCCGCAAGAAAGATACGAGATAGAGGAAACAGAAACAGTACCAAGTCTTACCAG 720
QY 965 AAGAAATGATTGGAGGATGCTGAAAGAAAAACAAGAAAGAAAGCAAGAAAAAGAGGAG 1024
DB 721 AAGAAATGATTGGAGGATGCTGAAAGAAAAACAAGAAAGAAAGCAAGAAAAAGAGGAG 780
QY 1025 GAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1084
DB 781 GAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 820
QY 1085 GAGAAAAACAACCTGAAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1144
DB 821 ----- 820
QY 1145 AGTTCAGAGAGCCTTAAACAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1204
DB 821 ----- 820
QY 1205 CATGAAAGAGATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1264
DB 821 ----- 820
QY 1265 GCAGAAAGAGAGAGAGAAATTAAGCCGACGACGACGACGACGACGACGACGACGACGAC 1324
DB 821 ----- 820
QY 1325 AGAATTGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1384
DB 821 ----- 820
QY 1385 AGGGAAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1444
DB 821 ----- 820
QY 1445 GAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1504
DB 821 ----- 820
QY 1505 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1564
DB 821 ----- 820
QY 1565 AGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1624
DB 821 ----- 820
QY 1625 GAGAAAAACATGCCATGCAAGAGACAAAGATATAAAGGGGAAAGAGTAGAACAGAAATA 1684
DB 821 ----- 820
QY 1685 GAAGGGAATGGGTAATGAAAAAGAAAGCACAAGAGATATAAATTCAGACAGCTGTCTTA 1744
DB 821 ----- 820
QY 1745 AAGAAACAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1804
DB 821 ----- 820
QY 1805 GACAAGCCTTCTTCAAAAAAGAAAGAGATCAAGATGAAAAAGATTAAGGAGCAAGAA 1864
DB 821 ----- 820
QY 1865 CCCAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1924
DB 856 CCCAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 915

```

```
2Y 1925 CAGAAATGAGAAATTCATGACCCACAACTTAAACATCTAGAAATCTTTTCAGCGCCCT 1984
2b 916 CAGAAATGAGAAATTCATGACCCACAACTTAAACATCTAGAAATCTTTTCAGCGCCCT 975
2Y 1985 GGAGGAGGCGCAGCGTGGACACCAAGAGAGCTGAGGGCGCCGCCACAGGTGGAAGCGCGC 2044
2b 976 GGAGGAGGCGCAGCGTGGACACCAAGAGAGCTGAGGGCGCCGCCACAGGTGGAAGCGCGC 1035
2Y 2045 AAAAGGCTGAGGAGCTTCGTCTGCTCGCGGGAGACCGAGAGCGAAGAGTTCGAGAAG 2104
2b 1036 AAAAGGCTGAGGAGCTTCGTCTGCTCGCGGGAGACCGAGAGCGAAGAGTTCGAGAAG 1095
2Y 2105 CTCAAAACAGAGACGACGAGCGCGCTTTGGAGCTGGAGAACTCAAGAAAAGAGCGAG 2164
2b 1096 CTCAAAACAGAGACGAGCGCGCTTTGGAGCTGGAGAACTCAAGAAAAGAGCGAG 1155
2Y 2165 GAGAGAAGGAAGCTTCGTGAGGAGGAAGAGCAGAGGAGGAAGCGAGGAGCGATCGA 2224
2b 1156 GAGAGAAGGAAGCTTCGTGAGGAGGAAGAGCAGAGGAGGAAGCGAGGAGCGATCGA 1215
2Y 2225 AAATCTCAGAGAGGAGAGAGAGGAGGCTTAAGAGAGAGATTGAAAGCGAGAGCA 2284
2b 1216 AAATCTCAGAGAGGAGAGAGAGGAGGCTTAAGAGAGAGATTGAAAGCGAGAGCA 1275
2Y 2285 GAAGCTGCTGAGAAACGCCAGAGATGCCAGAAAGATGGCTTGTTCAGATGACAAAGAACCA 2344
2b 1276 GAAGCTGCTGAGAAACGCCAGAGATGCCAGAAAGATGGCTTGTTCAGATGACAAAGAACCA 1335
2Y 2345 TTCAAGTGTTCACCTCTTAAAGTTTCATCTCTCAAGATAGAGAGCGAGCAATTTTGG 2404
2b 1336 TTCAAGTGTTCACCTCTTAAAGTTTCATCTCTCAAGATAGAGAGCGAGCAATTTTGG 1395
2Y 2405 AATAAGTGTGTGCAGAAAGCAGTGTGTCAATTCGACCCATCAAGCAGCAATAGTCTCC 2464
2b 1396 AATAAGTGTGTGCAGAAAGCAGTGTGTCAATTCGACCCATCAAGCAGCAATAGTCTCC 1455
2Y 2465 AAGATTGACAGAGACTGGAGCAGTATACAGTGTCAATTCAGGGACAAAAGCGCAAA 2524
2b 1456 AAGATTGACAGAGACTGGAGCAGTATACAGTGTCAATTCAGGGACAAAAGCGCAAA 1515
2Y 2525 CCTACAAAGCGCGAGCCTCGGATCTTCCTGTTCTCTGCTGAAAGTGTACCAACATCAAG 2584
2b 1516 CCTACAAAGCGCGAGCCTCGGATCTTCCTGTTCTCTGCTGAAAGTGTACCAACATCAAG 1575
2Y 2585 AGTATGTGGGAAAGGGAATGTGTTCATCCCCACTGACGAGGACACCAACAAAG 2644
2b 1576 AGTATGTGGGAAAGGGAATGTGTTCATCCCCACTGACGAGGACACCAACAAAG 1635
2Y 2645 GAAACTGTGCTTGAAGGTAGGGGTTCTAGCGGCATCAATGAATGGCTAACTAAACCC 2704
2b 1636 GAAACTGTGCTTGAAGGTAGGGGTTCTAGCGGCATCAATGAATGGCTAACTAAACCC 1695
2Y 2705 CCAGATGGAACCAAGTCACTGCTCCAAACCTTCTGACTTGAGACGAGAGGTATCC 2764
2b 1696 CCAGATGGAACCAAGTCACTGCTCCAAACCTTCTGACTTGAGACGAGAGGTATCC 1755
2Y 2765 AGCAAGCGGAACCTCTGGGAAAGCAATCTGTGGATAAGGTCACTTCCCCCACTAAGGTT 2824
2b 1756 AGCAAGCGGAACCTCTGGGAAAGCAATCTGTGGATAAGGTCACTTCCCCCACTAAGGTT 1815
2Y 2825 TGAGACAGTTCAGAAAGAACCAAGCTCAGAGCGGAGGACGAGCTCAGTGTAGAGGGC 2884
2b 1816 TGAGACAGTTCAGAAAGAACCAAGCTCAGAGCGGAGGACGAGCTCAGTGTAGAGGGC 1875
2Y 2885 TAATTCGCTCTGTTTGTATTTATGTTGATTTACTAAATTTGGGTTTCAATTTATTTT 2944
2b 1876 TAATTCGCTCTGTTTGTATTTATGTTGATTTACTAAATTTGGGTTTCAATTTATTTT 1935
2Y 2945 TTCAATATCCAGTAACCAAGTATATTAATCACTATATTAATTAATCAAGTCTAGAGA 3004
2b 1936 TTCAATATCCAGTAACCAAGTATATTAATCACTATATTAATTAATCAAGTCTAGAGA 1995
```

```
QY 3005 TGTTCATGTAAGAGTACTGCTCTTTCACAGAGAGCCTGTTTCTTAAAGAAACCCATGCTGT 3064
Db 1996 TGTTCATGTAAGAGTACTGCTCTTTCACAGAGAGCCTGTTTCTTAAAGAAACCCATGCTGT 2055
QY 3065 GAAATAGAGACTTTTCTACTGATCATCAATCTCTGATCTGAGCAGTGTGACCAACAC 3124
Db 2056 GAAATAGAGACTTTTCTACTGATCATCAATCTCTGATCTGAGCAGTGTGACCAACAC 2115
QY 3125 ATCTGAAAGTCAACAGAGAGATCCAAAGTTTAAATTTGCTGCGGAATGTGTGAGTATCTAG 3184
Db 2116 ATCTGAAAGTCAACAGAGAGATCCAAAGTTTAAATTTGCTGCGGAATGTGTGAGTATCTAG 2175
QY 3185 AAAATAGAACCTGCTGTTTGTGTTTAAATACAGAGATCATGTTGTTCTGCACTTTA 3244
Db 2176 AAAATAGAACCTGCTGTTTGTGTTTAAATACAGAGATCATGTTGTTCTGCACTTTA 2235
QY 3245 TAATAAGCATCGAAGAAATATCTTAGTAGGCAATTTGAACCTTTTGAAGTAACCC 3304
Db 2236 TAATAAGCATCGAAGAAATATCTTAGTAGGCAATTTGAACCTTTTGAAGTAACCC 2295
QY 3305 ATTTTCAGATTTCAATATCTGCGATTAATGTTGCTTTAAAAAAGAAAA 3354
Db 2296 ATTTTCAGATTTCAATATCTGCGATTAATGTTGCTTTAAAAAAGAAAA 2345

RESULT 11
LOCUS HUMLCAL 2198 bp mRNA linear PRI 18-DEC-2002
DEFINITION Homo sapiens mRNA for l-caldesmon 1, complete cds.
ACCESSION D90452
VERSION D90452.1 GI:219895
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Hayashi, K., Yano, H., Haehida, T., Takeuchi, R., Takeda, O., Asada, K.,
Takahashi, E., Kato, I. and Sobue, K.
TITLE Genomic structure of the human caldesmon gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (24), 12122-12126 (1992)
MEDLINE 93101679
PUBMED 1465449
REFERENCE 2 (bases 1 to 2198)
AUTHORS Hayashi, K. and Sobue, K.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-1991) Kenji Sobue, Osaka University, Medical
School, Biomedical Research Center, Department of Neurochemistry
and Neuroparmacology, Kita-ku, Osaka 530, Japan
(Tel:81-6-6445-6879, Fax:81-6-6445-6879)
FEATURES
source
1. .2198
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="Hela S3"
12..1688
/codon_start=1
/product="l-caldesmon 1"
/protein_id="BAAL14418.1"
/db_xref="GI:219896"
/trsnelation="MLGSGSHGRSLAALSOIAYORNDDEEBAERRRRARORERL
RQKEEESIGQVTDQVEVNAQNSVPDEEAKTTTTOVGEDEAAFLERLARERRRO
KRLQKALEKQKQKFDPTITDASLSPRRMQNDTAENETTEKEEKSESQERVEIETE
TVTKYQKNDWRDAENKKEDEEKEEKKPKKSGIGENGEKSGTKVQAKREKLQE
DKPTFKKEIKDEKIKKQKPEKVEKSMRKQKFTBVKSONGKFTMHLKHTEHTFS
RPGGASVDTKEAGAPQVAGKLELRELRGETESEEFEKQKQKQQAALAELEBLK
KRREERKVLLEEEQRKQKQADRLKLEERKEERKEEIERRRAAEAKRKQKMPDGL
SDPKPKCFTPKGSLKTEERAEKLVKSKGSKVSTHOAAITVSKIDSRLQVTSIA
IEGTSAKPTKPAASDLPVPAEGVRNPKSMWEGNVFSSPTAAGTPNKETAGLKVGVS
SRINELWTKTPGKNSPAPKPSDLRPGDVUSKKNLWEKQSVDKVTSPTKV"

CDS
ORIGIN
```


Query Match									
Best Local Similarity 42.1%; Score 1558; DB 9; Length 2198;									
Matches 1646; Conservative 94.6%; Pred. No. 0; Mismatches 90; Indels 4; Gaps 3;									
Qy	1601	CAGAAACGTAACAGCAGCTAGAGAGAGAAACATGCCATGCAAGAGACAAAGATAAAA	1660						
Db	462	CAAGAAAGATACAGATAGAGAGAAACAGAACAGTCCACCAAGAAATGAT	521						
Qy	1661	GGGGAAGAGTAAACAGAAATAGAGAGGAAATGGGTAAATGAAAGAAAGACCAAGAA	1720						
Db	522	TGGAGGATGCTGAGAGAAACAGAAAGAGAGACAGGAAAGAGGAGGAGAGAGAG	581						
Qy	1721	GATAACTTCAGACAGCTGCTTAAAGAAACAGGAGAGAGAGAGGAACTAAAGTGCA	1780						
Db	582	AAGCCAAAGCGAGGAGCATTTGGAGAAATCAGGAGAGAGAGAGGAACTAAAGTGCA	641						
Qy	1781	GCTAAAGAGAAAGCTCCAGAGAGACAGCCTACCTTCAAAAAAGAGAGATCAAGAT	1840						
Db	642	GCTAAAGAGAAAGCTCCAGAGAGACAGCCTACCTTCAAAAAAGAGAGATCAAGAT	701						
Qy	1841	GAAAGATTAAAG	1900						
Db	702	GAAAGATTAAAG	761						
Qy	1901	AAGGGATTACAGAGTTAAGTCCAGATCGAGATCGAGATCGAGATCGAGATCGAG	1960						
Db	762	AAGGGATTACAGAGTTAAGTCCAGATCGAGATCGAGATCGAGATCGAGATCGAG	821						
Qy	1961	ACTGAGATATCTTCAGCCGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2020						
Db	822	ACTGAGATATCTTCAGCCGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	881						
Qy	2021	GGGCCCCCAGGTGAG	2080						
Db	882	GGGCCCCCAGGTGAG	941						
Qy	2081	ACCGAGAGCGAAGAGTTCGAGAGAGCTCAACAGAGAGAGAGAGAGAGAGAGAG	2140						
Db	942	ACCGAGAGCGAAGAGTTCGAGAGAGCTCAACAGAGAGAGAGAGAGAGAGAGAG	1001						
Qy	2141	GAGGAATCAAGAAAG	2200						
Db	1002	GAGGAATCAAGAAAG	1061						
Qy	2201	AGGAAGCAGGAG	2260						
Db	1062	AGGAAGCAGGAG	1121						
Qy	2261	GAAGAGATTCAAG	2320						
Db	1122	GAAGAGATTCAAG	1181						
Qy	2321	GGTTGTGATGACAG	2380						
Db	1182	GGTTGTGATGACAG	1241						
Qy	2381	ATAGAAG	2440						
Db	1242	ATAGAAG	1301						
Qy	2441	ACCATCAAGCAG	2500						
Db	1302	ACCATCAAGCAG	1361						
Qy	2501	ATTGAGGAG	2560						
Db	1362	ATTGAGGAG	1421						
Qy	2561	GCTGAGGAG	2620						
Db	1422	GCTGAGGAG	1481						
Qy	2621	ACTGACAG	2680						

Db	1482	ACTGACAG	1541						
Qy	2681	ATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2740						
Db	1542	ATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1601						
Qy	2741	GACTTGAG	2800						
Db	1602	GACTTGAG	1661						
Qy	2801	AGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2860						
Db	1662	AGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1721						
Qy	2861	AGGACAG	2920						
Db	1722	AGGACAG	1779						
Qy	2921	AATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2980						
Db	1780	AATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1839						
Qy	2981	TATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3040						
Db	1840	TATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1898						
Qy	3041	TGTTTCTAAAG	3100						
Db	1899	TGTTTCTAAAG	1958						
Qy	3101	TATCTGAG	3160						
Db	1959	TATCTGAG	2018						
Qy	3161	CTGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3220						
Db	2019	CTGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2077						
Qy	3221	AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3280						
Db	2078	AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2137						
Qy	3281	TGTAACACTTTTGAAG	3340						
Db	2138	TGTAACACTTTTGAAG	2197						
RESULT 12									
HUMCALD									
Human caldesmon mRNA, complete cds.									
M64110.1 GI:179829									
P-actin binding protein; caldesmon; calmodulin; myosin; tropomyosin.									
Homo sapiens (human)									
SOURCE									
ORGANISM									
REFERENCE									
AUTHORS									
TITLE									
JOURNAL									
MEDLINE									
PUBMED									
COMMENT									
FEATURES									
source									
1..2975									
Location/Qualifiers									
/organism="Homo sapiens"									

2975 bp mRNA linear PRI 27-APR-1993

Human caldesmon mRNA, complete cds.

M64110.1 GI:179829

P-actin binding protein; caldesmon; calmodulin; myosin; tropomyosin.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2975)

Novy R.E., Lin J.L. and Lin J.J.

Characterization of cDNA clones encoding a human fibroblast caldesmon isoform and analysis of caldesmon expression in normal and transformed cells

J. Biol. Chem. 266 (25), 16917-16924 (1991)

91358497

1885618

Original source text: Homo sapiens (library: lambda gt11) fetus lung cDNA to mRNA.

Location/Qualifiers

1..2975

/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="Wi-38"
/cell_type="fibroblast"
/tissue_type="lung"
/dev_stage="fetus"
/tissue_lib="lambda gt11"
112..1728
/codon_start=1
/product="caldesmon"
/protein_id="Aa35636.1"
/db_xref="GI:179830"
/translators="MDFFERRRLRQKREEMLEBRIAYQVNDDEEBAERRRR
ACQRLRQKEESLEQVTDQVFNQNSVPDEAKTTNTQVGGDDDEAFLRLAR
EEROKRLOERLEQDFDPTTIDASLPSRQMDNTAENETKEESESQERY
EIEETETVTKSQKNDWABENKEDKEEKEEKPGRSIGENQIKOEKIKDKE
PKEEVKSFMDKKGFTEVKSQNGFMTLKHTEHTFSPGGRASVTKAEAGAPQME
AKRLEELRRRGETESEFELKQKQEALEELKKRERERKVLKEEEOQRKOE
EADRLREERERKRLKEEIERRAAEAKRQKMPEDGLSDDKPKPKFTPKGSLKIE
BRAEFLNKSVOKSGVKSTHQAAIVSKIDSLRLEQYTSAIESTKSAKTKPAASLPVP
AEGVRLIKSMKEGNVFSPTAAGTGNKETAGLKVGVSSRLNEWLTKTPDGNKSPAPK
PSDLRPGDVSSKRNLMKESVDKVTPTKV"

```

ORIGIN

```

Query Match      41.1%; Score 1522; DB 9; Length 2975;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1555; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1752 AGGGAAGAGAGAGGAACTAAAGTCAAGCTAAAGAGAGAAAGCTCCAGAGACAGC 1811
DB 653 AAGAAAACAAGAAAGAACACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 712
QY 1812 CTACCTTCAAAAAGAGATCAAGATGAAAGATTAAAGAGACAAAGAACCCAAAG 1871
DB 713 GGAGCATTTGGAGAAATCAGATCAAGATGAAAGATTAAAGAGACAAAGAACCCAAAG 772
QY 1872 AAGAAAGTTAAGAGCTTCATGGATCGAAGAGGAGATTACAGAAAGTTAAGTCGAGAGATG 1931
DB 773 AAGAAAGTTAAGAGCTTCATGGATCGAAGAGGAGGATTACAGAAAGTTAAGTCGAGAGATG 832
QY 1932 GAGAAATCATGACCAACAACTTAAACATACACTAGATATCTTTCAGCGCCCTCGAGGGA 1991
DB 833 GAGAAATCATGACCAACAACTTAAACATACACTAGATATCTTTCAGCGCCCTCGAGGGA 892
QY 1992 GGGCCAGCGTGACACCAAGAGGCTGAGGCGCGCCCGAGGAGGAGGAGGAGGAGGAGG 2051
DB 893 GGGCCAGCGTGACACCAAGAGGCTGAGGCGCGCCCGAGGAGGAGGAGGAGGAGGAGG 952
QY 2052 TGGAGGAGCTTCGTCTCGTCTCGCGGGAGACCGAGAGGAGAGGAGGAGGAGGAGGAGG 2111
DB 953 TGGAGGAGCTTCGTCTCGTCTCGCGGGAGACCGAGAGGAGGAGGAGGAGGAGGAGGAGG 1012
QY 2112 AGAAGCAGCAGGAGGCGGCTTTCGAGCTGGAGGAACTCAAGAAAAAGAGGAGGAGAGAA 2171
DB 1013 AGAAGCAGCAGGAGGCGGCTTTCGAGCTGGAGGAACTCAAGAAAAAGAGGAGGAGAGAA 1072
QY 2172 GGAAGTCTTGGAGGAGGAGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2231
DB 1073 GGAAGTCTTGGAGGAGGAGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1132
QY 2232 GAGAGGAGGAGAGAGAGAGGAGGCTTAAAGGAGAGATTGAAAGCGGAGAGAGAGAGGCTG 2291
DB 1133 GAGAGGAGGAGAGAGAGAGGAGGCTTAAAGGAGAGATTGAAAGCGGAGAGAGAGAGGAGGCTG 1192
QY 2292 CTGAGAAACCCAGAGAGATGCCAGAGAGATGGCTGTGTCAGATGACAAAGAACCATTCAGT 2351
DB 1193 CTGAGAAACCCAGAGAGATGCCAGAGAGATGGCTGTGTCAGATGACAAAGAACCATTCAGT 1252
QY 2352 GTTTCACCTCTAAAGTTCACTCTCAGATAGAGAGCGGAGCAATTTTGAATAAGT 2411
DB 1253 GTTTCACCTCTAAAGTTCACTCTCAGATAGAGAGCGGAGCGAGAAATTTTGAATAAGT 1312
QY 2412 CTGTGAGAGAAAGCAGTGTGTGTCATAATCGACCCATCAAGCAGCAATAGTCTCCAAAGATTG 2471

```

```

RESULT 13
AR001262
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

```

```

AR001262
Sequence 13 from patent US 5739008.
AR001262
AR001262.1
GI:3963329
Unknown.
SOURCE
Unknown.

```

```

2137 bp
DNA
linear
PAT 04-DEC-1998

```

1375	GGGAGAAAGGGAATGTGTTTTT	CATCCCCACTGCACGGCACACCAATAAGGAAACTG	1434
2652	CTGGCTTGAAGTAGGGTTC	TAGCCGCATCAATGAATGGCTAACTAAACCCCCAGATG	2711
1435	CTGGCTTGAAGTAGGGTTC	TAGCCGCATCAATGAATGGCTAACTAAACCCCCAGATG	1494
2712	GAAACAAGTCACTGCTGCC	CCAAACCTTCTGACTTGAGACCGAGGAGCGTATCCAGCAAGC	2771
1495	GAAACAAGTCACTGCTGCC	CCAAACCTTCTGACTTGAGACCGAGGAGCGTATCCAGCAAGC	1554
2772	GGAACTCTGGGAAGCAATCT	GTGGGATAAGGTCACTTCCCCCACTTAAGGGTTTGAGACA	2831
1555	GGAACTCTGGGAAGCAATCT	GTGGGATAAGGTCACTTCCCCCACTTAAGGGTTTGAGACA	1614
2832	GTTCAGAAAGAACCCAACT	CAAGCGCAGGACGAGCTCAGTTGTAGAGGGCTAATTCG	2891
1615	GTTCAGAAAGAACCCAACT	CAAGCGCAGGACGAGCTCAGTTGTAGAGGGCTAATTCG	1674
2892	CTCTGTTTGTATTTATGTT	GATTACTAAATTTGGGTTCAATATCTTTTATTTTCAATA	2951
1675	CTC--TTTGTATTATGTTG	ATTACTAAATTTGGGTTCAATATCTTTTATTTTCAATA	1732
2952	TCCCAGTAAACCCATGATAT	TATCACTATATTTAATAATCAAGCTCTAGAGATGTTTCAT	3011
1733	TCCCAGTAAACCCATGATAT	TATCACTATATTTAATAATCAAG--CTAGAGATGTTTCAT	1791
3012	GGTAAAGTACTGCCCTTTC	CACAGGAGCGCTGTTCTTAAAGAAACCCCATGCTGTGAAATAG	3071
1792	GGTAAAGTACTGCCCTTTC	CACAGGAGCGCTGTTCTTAAAGAAACCCCATGCTGTGAAATAG	1851
3072	AGACTTTTCTACTGATCAT	CACTCTGATCTGAGCAGTGATACCAACCATCTGAA	3131
1852	AGACTTTTCTACTGATCAT	CACTCTGATCTGAGCAGTGATACCAACCATCTGAA	1911
3132	GTCAACAGAAAGTCCAAAG	TTTAAATTTGCCTGGGGAATGTGTCAGTATCTAGAAAAATG	3191
1912	GTCAACAGAAAGTCCAAAG	TTTAAATTTGCCTGGGGAATGTGTCAGTATCTAGAAAAATG	1971
3192	AACCGTAGTTTTGTTTTTT	TAAATCAGAAAGTCATGTTGTTCTGCACATTATAATAAA	3251
1972	AACCGTAG--TTTTGTTTT	TAAATCAGAAAGTCATGTTGTTCTGCACATTATAATAAA	2030
3252	GCATGGAAGAAATATCTCT	TAGTAGGCAATTTGAACACTTTTTTGAAGTAACCCATTTTCAG	3311
2031	GCATGGAAGAAATATCTCT	TAGTAGGCAATTTGAACACTTTTTTGAAGTAACCCATTTTCAG	2090
3312	ATTTGAAATCTCGATATGTT	TGCTTTTAAAAAAGAAAAAGAAAGA	3357
2091	ATTTGAAATCTCGAATAAT	GGTTGTTCTTTAAAAAAGAAAAA	2136
RESULT 14			
123420		2137 bp	DNA
LOCUS			linear
DEFINITION	Sequence 13 from patent US 5532337.		
ACCESSION	123420		
VERSION	123420.1		
KEYWORDS	GI:1603290		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 2137)		
AUTHORS	Hayashi,K., Hashida,T., Asada,K., Kotani,H., Kato,I. and Sobue,K.		
TITLE	Proteins comprising calmodulin- and actin-binding human caldesmon		
JOURNAL	polypeptide fragments		
FEATURES	Patent: US 5532337-A 13 02-JUL-1996;		
source	Location/Qualifiers		
	1..2137		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			

Db	1555	GGAACTCTCTGGGAAAAGCAATCTGTGGATAAGTCTCACTTCCCCACCTAAGGTTTGAGACA	1614
Qy	2832	GTTCGAGAAAGAACCCAGAGCTCAAGACGCGAGGACGAGCTCAGTTGTAGAGGGCTAATTCG	2891
Db	1615	GTTCGAGAAAGAACCCAGAGCTCAAGACGCGAGGACGAGCTCAGTTGTAGAGGGCTAATTCG	1674
Qy	2892	CTCTGTTTTTGTATTTATGTGGTATTACTAAATTTGGGTTTCATTAATCTTTTATTTTCAATA	2951
Db	1675	CTC--TTTGTATTTATGTGGTATTACTAAATTTGGGTTTCATTAATCTTTTATTTTCAATA	1732
Qy	2952	TCCAGTAAACCCATGTATATTAATCACTATTAATTAATACAGTCTAGAGATGTTTCAT	3011
Db	1733	TCCAGTAAACCCATGTATATTAATCACTATTAATTAATTAATCAAG-CTAGAGATGTTTCAT	1791
Qy	3012	GGTAAAGTACTGCTTTTGACAGGAGCCTGTTTCTAAAGAAACCCCATGCTGTGAAATAG	3071
Db	1792	GGTAAAGTACTGCTTTTGACAGGAGCCTGTTTCTAAAGAAACCCCATGCTGTGAAATAG	1851
Qy	3072	AGACTTTTCTACTGATCATCATTAACCTGTGTATCTGAGCAGTGATACCAACACATCTGAA	3131
Db	1852	AGACTTTTCTACTGATCATCATTAACCTGTGTATCTGAGCAGTGATACCAACACATCTGAA	1911
Qy	3132	GTCAACAGAAAGATCCAAAGTTTAAATTTGCCTGCGGAAATGTGTGCAGTATCTAGAAAAATG	3191
Db	1912	GTCAACAGAAAGATCCAAAGTTTAAATTTGCCTGCGGAAATGTGTGCAGTATCTAGAAAAATG	1971
Qy	3192	AACCGTAGTTTTTGTGTTTTTAAATACAGAAAGTCATGCTTCTGTCACATTTATAATAAA	3251
Db	1972	AACCGTAG--TTTTGTGTTTTTAAATACAGAAAGTCATGCTTCTGTCACATTTATAATAAA	2030
Qy	3252	GCATGGAAGAAATTTATCTTAGTAGGCAATTTGAACACTTTTGAAGTAAACCCATTTTCAG	3311
Db	2031	GCATGGAAGAAATTTATCTTAGTAGGCAATTTGAACACTTTTGAAGTAAACCCATTTTCAG	2090
Qy	3312	ATTGTGAATCTGCGAATAGTGGTTGCTTTTAAAGAAAGAAAAAGA	3357
Db	2091	ATTGTGAATCTGCGAATAGTGGTTGCTTTTAAAGAAAGAAAAAGA	2136
RESULT 15			
HUMLCA2			
LOCUS	HUMLCA2	2120 bp	mRNA linear PRI 18-DEC-2000
DEFINITION	Homo sapiens mRNA for l-caldesmon II, complete cds.		
ACCESSION	D90453		
VERSION	D90453.1	GI:219897	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Hayashi, K., Yano, H., Hashida, T., Takeuchi, R., Takeda, O., Asada, K., Takahashi, E., Kato, I. and Sobue, K.		
TITLE	Genomic structure of the human caldesmon gene		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 89 (24), 12122-12126 (1992)		
MEDLINE	93101679		
PUBMED	1465449		
REFERENCE	2	(bases 1 to 2120)	
AUTHORS	Hayashi, K. and Sobue, K.		
TITLE	Direct Submision		
JOURNAL	Submitted (26-JUL-1991) Kenji Sobue, Osaka University, Medical School, Biomedical Research Center, Department of Neurochemistry and Neuroparmacology, Kita-Ku, Osaka 530, Japan (Tel:81-6-6445-6879, Fax:81-6-6445-6879)		
FEATURES	Location/Qualifiers		
source	1..2120		/organism="Homo sapiens"
			/mol_type="mRNA"
			/db_xref="taxon:9606"
			/cell_line="HeLa S3"
			12..1610
			/codon_start=1
CDS			

```

/product="1-caldesmon I"
/protein_id="BAAI4419.1"
/db_xref="GI:219898"
/translation="*WLGGSGHGRSLAALSQLVQRNDDDEEAERRRRPARQERL
RQREERLGGVDTQVEVGNASVPDEAKTNTTNVQEGDDAAFLERLARERERQ
KLQEAALERQKEFTPTIDASLVSFRRMQNDTANETTEEKESQREYIEBTE
TVTKYQKNDWRDAENKDEKDEEEEEKPKRIGENQIKDKIKKDEPKREVK
SFMDRKGFVTSGKQGVFMFKLTKHTNTSFPGRAVSVTKEAGAPVQVAGRKLE
ELRERGETESEEKPKQKQEALELELKKREERKVLREEQSRQKQEAARDKL
REERKRLKZEIERRAAEAKQKQMPEDGLDDKKPKCTPKGSSLKIEERAEFL
NKSVQSGSKVTHQAATVSKIDSLRQETVAIGETKAKPTKPAASDLFVPAEGVRN
IKSMWEGKNSPSSTAQTPNKETAGLKVGVYSRINELTKTPDNGKSGPAPKPSDLRP
GVNASRNIWPKQGVKVTSPKV"

```

ORIGIN

Query Match	39.7%	Score 1469.4	DB 9	Length 2120
Best Local Similarity	96.5%	Pred. No. 6.6e-304		
Matches 1534	Conservative	0	Mismatches 51	Indels 4
Gaps	3			
QY	1752	AGGGAGAAGAGAGGGAACTTAAAGTGCAGACTAAAAGAGAAAAAGCTCCAAGAGACAAG	1811	
DB	535	AAGAAACACAGAAAGAACACAGGAAAGGAGGAGGAGGAAGAGGAAAGCCAAAGCGAG	594	
QY	1812	CTACCTTCAAAAAGAGAGATCAAGATCAAAAGATTAAAAAGGACAAAGACCCCAAG	1871	
DB	595	GGAGCATTTGAGAAAAATCAGATCAAAAGATGAAAAGATTAAAAAGGACAAAGAACCCCAAG	654	
QY	1872	AAGAAGTTAAGAGCTTCATCGATCGAAGAGGGATTTACAGAAGTTTAAAGTCGCAATG	1931	
DB	655	AAGAAGTTAAGAGCTTCATCGATCGAAGAGGGATTTACAGAAGTTTAAAGTCGCAATG	714	
QY	1932	GAGAAATCATGACCCCAAACTTTAAACATACTCAGAAATACTTTTCAGCGCCCTCGAGGGA	1991	
DB	715	GAGAAATCATGACCCCAAACTTTAAACATACTCAGAAATACTTTTCAGCGCCCTCGAGGGA	774	
QY	1992	GGGCCAGCGTGGACACCAAGGAGCTGAGGCGCCCCCAGGTGGAGCGCGCAAAAGGC	2051	
DB	775	GGGCCAGCGTGGACACCAAGGAGCTGAGGCGCCCCCAGGTGGAGCGCGCAAAAGGC	834	
QY	2052	TGAGAGAGCTTCGTGCTGCTCGGGGAGACCCAGAGCGAAGAGTTTCGAGAAGCTCAAAC	2111	
DB	835	TGAGAGAGCTTCGTGCTGCTCGGGGAGACCCAGAGCGAAGAGTTTCGAGAAGCTCAAAC	894	
QY	2112	AGAAGCAGCAGGAGGCGGCTTTTGAGCTGAGGAACTCAAGAAAAAGAGGGAGGAGAGAA	2171	
DB	895	AGAAGCAGCAGGAGGCGGCTTTTGAGCTGAGGAACTCAAGAAAAAGAGGGAGGAGAGAA	954	
QY	2172	GGAAGTCTCTGGAGGAGGAAGACAGAGGAGGAGAGCAGGAGGAGCGATCGGAAAACTCA	2231	
DB	955	GGAAGTCTCTGGAGGAGGAAGACAGAGGAGGAGAGCAGGAGGAGCGATCGGAAAACTCA	1014	
QY	2232	GAGAGGAGGAGAGAGAGGAGGCTTAAGGAGAGATTTGAAAGCGGAGAGCAGAGAGCTG	2291	
DB	1015	GAGAGGAGGAGAGAGAGGAGGCTTAAGGAGAGATTTGAAAGCGGAGAGCAGAGAGCTG	1074	
QY	2292	CTGAGAAACGCGCAGAGATGCGCAAGATGCGCTTGTCAGATGACAAAGAAACCAATCAAAGT	2351	
DB	1075	CTGAGAAACGCGCAGAGATGCGCAAGATGCGCTTGTCAGATGACAAAGAAACCAATCAAAGT	1134	
QY	2352	GTTTCACTCCTAAAGGTTTCATCTCTCAAGATAGAAAGCGAGCGAGAAATTTTGAATAGT	2411	
DB	1135	GTTTCACTCCTAAAGGTTTCATCTCTCAAGATAGAAAGCGAGCGAGAAATTTTGAATAGT	1194	
QY	2412	CTGTGCAGAAAAAGCAGTGGTGTCAAAATCGACCCATCAAGCAGCAATAGTCTCCAAAGTTG	2471	
DB	1195	CTGTGCAGAAAAAGCAGTGGTGTCAAAATCGACCCATCAAGCAGCAATAGTCTCCAAAGTTG	1254	
QY	2472	ACAGCAGCTGGAGCAGTATACCAATTTGAGGGAACAAAAGCGCAAAACCTTCAAA	2531	
DB	1255	ACAGCAGCTGGAGCAGTATACCAATTTGAGGGAACAAAAGCGCAAAACCTTCAAA	1314	
QY	2532	AGCCGGCAGCCTCGGATCTTCCTGTTCTGCTGAAGTGTGACGCAATCAAGATATGT	2591	

Search completed: March 11, 2004, 03:36:40
Job time : 9450.45 secs

1315	AGCGGCAGCCTCGGATCTTCTGTCTCTGTCTGAAGGTGTACGCCACATCAAGATATGT	1371
2592	GGGAGAAAGGGAATGTGTCTTATCCCCCACTGACAGAGGCACACCAATTAAGGAAACTG	2651
1375	GGGAGAAAGGGAATGTGTCTTATCCCCCACTGACAGAGGCACACCAAAATGAGGAACTG	1434
2652	CTGGCTTGAAGGTAGGGGTCTTAGCCGCATCAATGAATGGCTTAATAAAACCCCAAGATG	2711
1435	CTGGCTTGAAGGTAGGGGTCTTAGCCGTATCAATGAATGGCTTAATAAAACCCCAAGATG	1494
2712	GAAACAAGTCACTGCTGCCAAACCTTTCTGACCTTGAGACCGAGAGAGGTTATCCAGCAAGC	2771
1495	GAAACAAGTCACTGCTGCCAAACCTTTCTGACTTGGAGCCAGAGAGCGTATCCAGCAAGC	1554
2772	GGAACTCTGGGAAAAGCAATCTGTGGATAAGGTCACTTCCCCCACTAAGGTTTGAGACA	2831
1555	GGAACTCTGGGAAAAGCAATCTGTGGATAAGGTCACTTCCCCCACTAAGGTTTGAGACA	1614
2832	GTTCCAGAAAGAACCCAAAGCTCAAGACGGCAGGACGAGCTCAGTTGTAGGGGCTAATTCG	2891
1615	GTTCCAGAAAGAACCCAAAGCTCAAGACGGCAGGACGAGCTCAGTTGTAGGGGCTAATTCG	1674
2892	CTCTGTTTTGTAATTTATGTGTGTTAACTTAAATTTGGGTTCATTAATCTTTTATTTTCAATA	2951
1675	CTC--TTTTGTATTTATGTGTGTTAACTTAAATTTGGGTTCATTAATCTTTTATTTTCAATA	1732
2952	TCCAGTAACCCCAATATTAATCACTATATTTAATAATCAAGTCTAGAGATGTTTCAT	3011
1733	TCCAGTAACCCCAATATTAATCACTATATTTAATAATCAAG--CTAGAGATGTTTCAT	1791
3012	GGTAAAGTACTGCCCTTTGCACAGGAGCCTGTTTCTAAAGAAACCCCAATGCTGTGAATAAG	3071
1792	GGTAAAGTACTGCCCTTTGCACAGGAGCCTGTTTCTAAAGAAACCCCAATGCTGTGAATAAG	1851
3072	AGACTTTCTACTGTATCATCATTAACCTCTGTATCTGAGCAGTGATACCAACCATCTGAA	3131
1852	AGACTTTCTACTGTATCATCATTAACCTCTGTATCTGAGCAGTGATACCAACCCATCTGAA	1911
3132	GTCAACAGAGATCCAAAGTTTAAATTTGCCCTGGGAATGTGTCAGTATCTAGAAAAATG	3191
1912	GTCAACAGAGATCCAAAGTTTAAATTTGCCCTGGGAATGTGTCAGTATCTAGAAAAATG	1971
3192	AACCGTAGTTTTTGTTTTTTAAATACAGAAATCATGTTCTTTCTGCACTTTATATAATAA	3251
1972	AACCGTAG--TTTTTGTTTTTTAAATACAGAAATCATGTTCTTTCTGCACTTTATATAATAA	2030
3252	GCATGAGAGAAATTAATCTAGTAGGCAATGTTAACAATTTTGAAGAGTAACCCATTTTCAG	3311
2031	GCATGAGAGAAATTAATCTAGTAGGCAATGTTAACAATTTTGAAGAGTAACCCATTTTCAG	2090
3312	ATTTGAATACTCGGATAATGGTTTGTCTTT	3340
2091	ATTTGAATACTCGAATAATGGTTGTCTT	2119

Search completed: March 11, 2004, 03:36:40
Job time : 9450.45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 08:01:57 ; Search time 900.994 Seconds
(without alignments)
17454.974 Million cell updates/sec

Title: US-10-084-817-30

Perfect score: 3702

Sequence: 1 gcctaggaaatgacaggcat.....aaggggcggttaccgagaac 3702

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124039041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseqn_29Jan04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002s:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3116.4	84.2	3610	6	ABL62384 Colon ade
2	3116.4	84.2	3610	6	ABL68604 Kidney ca
3	3116.4	84.2	3610	6	ABN95909 Gene #240
4	1793.6	48.4	4547	7	ABX77534 Different
5	1567	42.3	2215	2	AAQ28988 Sequence
6	1455.8	39.3	3342	4	AAH23102 Osteoarthritis
7	1452.8	39.2	2137	2	AAQ28987 Sequence
8	1086.2	29.3	1677	2	AAQ28986 Sequence
9	999.2	27.0	1599	2	AAQ28985 Sequence
10	932.6	25.2	539	2	AAQ28984 Sequence
11	691	18.7	705	3	AAC00905 Human sec
12	689.2	18.6	756	2	AAQ39869 Gastric c
13	597.6	16.1	706	2	AAQ39870 Gastric c
14	584.2	15.8	2511	3	AAQ39870 Human can
15	584.2	15.8	2511	4	AAH33117 Human col
16	546.8	14.8	2209	7	ABX77533 Different
17	546.8	14.7	548	3	AAC01571 Human sec
18	486	13.1	1750	2	AAQ23748 1-Caldesmon
19	479.8	13.0	1014	2	AAQ12340 Caldesmon
20	447.4	12.1	449	6	ABV95786 Human pan
21	436.6	11.8	503	8	ACH32073 Human end
22	428.6	11.6	436	8	ACH15816 Human adu
23	428.6	11.6	606	3	AAAL6290 Human col

C

24	392	10.6	604	5	ABV55814 Human pro
25	377.2	10.2	1004	6	ABN74587 Bovine em
26	350.6	9.5	422	4	AAI90712 Human pol
27	349.2	9.4	369	2	AAQ28983 Sequence
28	348.8	9.4	357	2	AAQ28982 Sequence
29	346.4	9.4	348	2	AAQ28981 Sequence
30	331.4	9.0	504	4	AAI25076 Human bre
31	284	7.7	292	6	ABL75838 Corn tass
32	264.2	7.1	273	2	AAQ28998 Sequence
33	263	7.1	285	2	AAQ28996 Sequence
34	229.6	6.2	339	3	AAQ78295 Human can
35	228.4	6.2	492	3	ABQ55690 Human ova
36	211.4	5.7	400	7	ABX38333 Bovine ES
37	191	5.2	3662	4	ABL05274 Drosophila
38	189	5.1	427	3	AAC10424 Human sec
39	170	4.6	3399	2	AAT05868 Chicken 1
40	168	4.5	1050	4	AAQ21685 Mutational
41	168	4.5	2849	4	AAQ21684 Human ret
42	167	4.5	3579	3	AAA70099 Plasmodiu
43	164.2	4.4	1840	9	ADC87550 Human GPC
44	162.4	4.4	618	6	ABN65775 Human can
45	149.4	4.0	1662	4	ABL05275 Drosophila

ALIGNMENTS

RESULT 1

ABL62384

ID ABL62384 standard; DNA; 3610 BP.

XX ABL62384;

AC ABL62384;

XX 15-MAY-2002 (first entry)

DT 15-MAY-2002 (first entry)

XX 15-MAY-2002 (first entry)

XX Colon adenocarcinoma related gene sequence SEQ ID NO:721.

DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

XX Stomach; lung; prostate; pancreas; carcinoma; antitumor; carcinosarcoma;

XX Cystostatic; Gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

XX Gene; ds.

XX Homo sapiens.

OS Homo sapiens.

XX WO200194629-A2.

PN WO200194629-A2.

XX 13-DEC-2001.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US010838.

XX 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

XX 05-JUN-2000; 2000US-0209531P.

XX 18-SEP-2000; 2000US-0233133P.

XX 18-SEP-2000; 2000US-0233617P.

XX 20-SEP-2000; 2000US-0234009P.

XX 20-SEP-2000; 2000US-0234034P.

XX 20-SEP-2000; 2000US-0234052P.

XX 22-SEP-2000; 2000US-0234509P.

XX 22-SEP-2000; 2000US-0234567P.

XX 25-SEP-2000; 2000US-0234923P.

XX 25-SEP-2000; 2000US-0235082P.

XX 25-SEP-2000; 2000US-0235134P.

XX 25-SEP-2000; 2000US-0235280P.

XX 26-SEP-2000; 2000US-0235637P.

XX 26-SEP-2000; 2000US-0235638P.

XX 27-SEP-2000; 2000US-0235711P.

XX 27-SEP-2000; 2000US-0235720P.

XX 27-SEP-2000; 2000US-0235840P.

XX 27-SEP-2000; 2000US-0235863P.

XX 28-SEP-2000; 2000US-0236028P.

XX 28-SEP-2000; 2000US-0236032P.

1321 AAGGCGCAGCAGAGGAGGCGCAAGGCGCCAGGCGCAGAGGAGGAGAGGCTTAAGGTAGA 1380
1597 AGACGAGAAAAGCTTAACAGCAGCTAGAGAGAGAAAAACATGCGCATGCCAAGAGAGCAAAAGAT 1656
1381 AGACGAGAAAAGCTTAACAGCAGCTAGAGAGAGAAAAACGTCGCCATCAAGAGAGCAAAAGAT 1440
1557 AAAAGGGGAAAAGGTAGAACACAGAAAATAGAGGGGAANTGGTAAATGAAAAGAACGACAA 1716
1441 AAAAGGGGAAAAGGTAGAACACAGAAAATAGAGGGGAANTGGTAAATGAAAAGAACGACAA 1500
1717 AGAAGATAAACTTCAGACAGAGCTGTCTTAAGAGAAACAGAGGAGAGAGAGAGGGAATAAAGT 1776
1501 AGAAGATAAACTTCAGACAGAGCTGTCTTAAGAGAAACAGAGGAGAGAGAGAGGGAATAAAGT 1560
1777 GCAAGCTTAAAGAGAAAAGCTTCAAGAGACAAAGCCTACCTTCAAAAAGAGAGAGATCAA 1836
1561 GCAAGCTTAAAGAGAAAAGCTTCAAGAGACAAAGCCTACCTTCAAAAAGAGAGAGATCAA 1620
1837 AGATGAAAAGATTAAAAAGGACAAAGAACCCAAAGAGAAAGTTAAGAGCTTTCATGGATCG 1896
1621 AGATGAAAAGATTAAAAAGGACAAAGAACCCAAAGAGAAAGTTAAGAGCTTTCATGGATCG 1680
1897 AAAGAGGGATTTACAGAAAGTTAAGTCGAGAAATGGAGAAATTCATGACCCACAAACTTAA 1956
1681 AAAGAGGGATTTACAGAAAGTTAAGTCGAGAAATGGAGAAATTCATGACCCACAAACTTAA 1740
1957 ACATACTCAGAAATACCTTTTCAGCGCCCTCGAGGGAGGCGCAGCGTGGACACCAAGGAGGC 2016
1741 ACATACTCAGAAATACCTTTTCAGCGCCCTCGAGGGAGGCGCAGCGTGGACACCAAGGAGGC 1800
2017 TGAGGGGCCCCCGCCAGGTGGAAAGCCGGCAAAAAGGCTGGAGGAGCTTCGTGCTGTCGCGG 2076
1801 TGAGGGGCCCCCGCCAGGTGGAAAGCCGGCAAAAAGGCTGGAGGAGCTTCGTGCTGTCGCGG 1860
2077 GGAGACCGAGAGCAGAGAGCTTCGAGAAAGCTCAACAGAGACAGCAGGAGCGGCTTTGGA 2136
1861 GGAGACCGAGAGCAGAGAGCTTCGAGAAAGCTCAACAGAGACAGCAGGAGCGGCTTTGGA 1920
2137 GCTGGAGAACTCAAGAAAAGAGGAGGAGAGAGAAAGGCTCCTGGAGGAGAGAGCA 2196
1921 GCTGGAGAACTCAAGAAAAGAGGAGGAGAGAGAAAGGCTCCTGGAGGAGAGAGCA 1980
2197 GAGGAGGAGCAGGAGGAGGAGCGCATCGAAAACCTCAGAGAGGAGGAGAGAGAGAGGAGGCT 2256
1981 GAGGAGGAGCAGGAGGAGGAGCGCATCGAAAACCTCAGAGAGGAGGAGAGAGAGAGGAGGCT 2040
2257 AAAGGAGAGAGATTGAAAGCGGAGAGCAGAGAGCTGCTGAGAAAACGCCAGAGAGATGCCAGA 2316
2041 AAAGGAGAGAGATTGAAAGCGGAGAGAGCAGAGAGCTGCTGAGAAAACGCCAGAGAGATGCCAGA 2100
2317 AGATGGCTGTGATGACAGAGAAACCATTCAGAGTGTTCCTCACTCCCTAAAGGTTTCATCTCT 2376
2101 AGATGGCTGTGATGACAGAGAAACCATTCAGAGTGTTCCTCACTCCCTAAAGGTTTCATCTCT 2160
2377 CAAGATAGAGAGCGAGCAGAGATTTTGAATAAGTCTGTGAGAAAAGAGCAGTGCTGTCAA 2436
2161 CAAGATAGAGAGCGAGCAGAGATTTTGAATAAGTCTGTGAGAAAAGAGCAGTGCTGTCAA 2220
2437 ATCGACCCATCAGAGCAGATAGCTTCCAGAGTTGACAGCAGACTGGAGCAGATATACAG 2496
2221 ATCGACCCATCAGAGCAGATAGCTTCCAGAGTTGACAGCAGACTGGAGCAGATATACAG 2280
2497 TGCAATTGAGGGAAACAAAAGCGCAAAACCTTACAAAGCCGGCAGCCTCGGATCTTCCTGT 2556
2281 TGCAATTGAGGGAAACAAAAGCGCAAAACCTTACAAAGCCGGCAGCCTCGGATCTTCCTGT 2340
2557 TCCTGCTGAGAGTGTACGACACATCAAGAGTATGTGGAGAAAGGGAATGTGTTTTCATC 2616
2341 TCCTGCTGAGAGTGTACGACACATCAAGAGTATGTGGAGAAAGGGAATGTGTTTTCATC 2400
2617 CCCCCTGAGAGGAGGAGCAGCAAAATTAAGGAAACCTGCTGGCTTGAAGGTAGGGGTTTCTAG 2676

2401 CCCACTGCGAGCGCACACCAAATAAGGAARACTGCTGGCTTGAAGTAGGGGTTTCTAG 2460
2677 CCGCATCAATGAATGGCTAACTAAACCCAGATGGAACCAAGTCACTGCTCCCAAAACC 2736
2461 CCGCATCAATGAATGGCTAACTAAACCCAGATGGAACCAAGTCACTGCTCCCAAAACC 2520
2737 TTCTGACTTGAGACCCAGAGAGCTATCCAGCAAGCGGAACCTCTGGGAAAAGCAATCTGT 2796
2521 TTCTGACTTGAGACCCAGAGAGCTATCCAGCAAGCGGAACCTCTGGGAAAAGCAATCTGT 2580
2797 GGATAAGTCTACTTCCCCACATAAGGTTTGAGACAGTTCCAGAAAAGCAACCAAGCTCAAG 2856
2581 GGATAAGTCTACTTCCCCACATAAGGTTTGAGACAGTTCCAGAAAAGCAACCAAGCTCAAG 2640
2857 ACGCAGGACGAGCTCAGTTGTAGAGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTT 2916
2641 ACGCAGGACGAGCTCAGTTGTAGAGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTT 2700
2917 ACTAAATGGGTTCAATATCTTTTATTTTCAATATCCCAAGTAAACCCATGTATATTATC 2976
2701 ACTAAATGGGTTCAATATCTTTTATTTTCAATATCCCAAGTAAACCCATGTATATTATC 2760
2977 ACTATATTTAATAATCACAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTCACACAGG 3036
2761 ACTATATTTAATAATCACAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTCACACAGG 2820
3037 AGCCTGTTTCTAAAGAAAACCCATGCTGTGAAATGAGAGACTTTTCTACTGATCATATAAC 3096
2821 ATCCTGTTTCTAAAGAAAACCCATGCTGTGAAATGAGAGACTTTTCTACTGATCATATAAC 2880
3097 TCTGATCTGAGCAGTGATACCAACCCATGCTGAAAGTCAACAGAGATCCAAGTTTAAAA 3156
2881 TCTGATCTGAGCAGTGATACCAACCCATGCTGAAAGTCAACAGAGATCCAAGTTTAAAA 2940
3157 TTGCTCGGGAATGTGTCAGTATCTAGAAAAGTAAACCCGATGTTTGTGTTTAAAAAT 3216
2941 TTG-CTGCGGAATGTGTCAGTATCTAGAAAAGTAAACCCGATGTTTGTGTTTAAAAAT 2999
3217 ACAGAGTCACTGTTGTTTCTGCACTTTTATAATAAGCATGGAAGAAATTTATCTAGTAGG 3276
3000 ACAGAGTCACTGTTGTTTCTGCACTTTTATAATAAGCATGGAAGAAATTTATCTAGTAGG 3059
3277 CAATTGTAACACTTTTGAAGTAACCCATTTTCAGATTTTGAATACTGCAATATCTAGTAGG 3336
3060 CAATTGTAACACTTTTGAAGTAACCCATTTTCAGATTTTGAATACTGCAATATCTAGTAGG 3119
3337 CTTTAAAAAAAAGAAAAGATGCGCGGACAGTTATTCCTTT 3378
3120 TCTTTAAAAAAAAGAAAAGATGCTGTTAGGTATTACTTTT 3161

RESULT 2

ABL68604

ID ABL68604 standard; DNA; 3610 BP.

XX ABL68604;

XX 15-MAY-2002 (first entry)

XX Kidney cancer related gene sequence SEQ ID NO:6941.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

XX cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;

XX Gene; ds.

XX Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US010838.

1b	961	GGAAACAAGGTT	CAGATGAGATT	TTCCCATCAT	GAAAAGATTG	CAAGAGGAAGA	CAAGGAAG	1020
1y	1237	AGCTGAGGCAGAGGGCA	AGGTTGGAGGACGA	GAAGAAAGAGAAAGAA	TTAAAGCCGAGCA	1296		
1b	1021	AGCTGAGGCAGAGGGCA	AGGTTGGAGGACGA	GAAGAAAGAGAAAGAA	TTAAAGCCGAGCA	1080		
1y	1297	AGACAAAAGATAGCAGAT	GAAACGACGAAGAA	TTGAAGCAGAGAAAGAA	AAAGACAGCTGCACA	1356		
1b	1081	AGACAAAAGATAGCAGAT	GAAACGACGAAGAA	TTGAAGCAGAGAAAGAA	AAAGACAGCTGCACA	1140		
1y	1357	AGAAAGAGAAAGGACAGAG	GGCAGAGAGGCAAGGA	ATGAGGAGGAGAGCAAGAA	AGGAGCAAGAAAGGC	1416		
1b	1141	AGAAAGAGAAAGGACAGAG	GGCAGAGAGGCAAGGA	ATGAGGAGGAGAGCAAGAA	AGGAGCAAGAAAGGC	1200		
1y	1417	AGCAGAGGAGGACAGAGAT	TAAGGAGGAGAAAGGG	CAGCAGAGGACAGGACAG	AG 1476			
1b	1201	AGCAGAGGAGGACAGAGAT	TAAGGAGGAGAAAGGG	CAGCAGAGGACAGGACAG	AG 1260			
1y	1477	GATAAAGAGGAGAGAAAG	GGCAGCAGAGAGAGG	CAGAGGACAGAGATTA	AAAGAGGAGAGAA 1536			
1b	1261	GATAAAGAGGAGAGAAAG	GGCAGCAGAGAGAGG	CAGAGGACAGAGATTA	AAAGAGGAGAGAA 1320			
1y	1537	AAGGGCAGCAGAGGAGAG	CAAGCAAGGGCCAGGGC	CAGAGGAGGAGAGAGGCT	TAAGGTAGA 1596			
1b	1321	AAGGGCAGCAGAGGAGAG	CAAGCAAGGGCCAGGGC	CAGAGGAGGAGAGAGGCT	TAAGGTAGA 1380			
1y	1597	AGAGCAGAAAACGTAA	CAAGCAGCTTAGAGAG	AAAAAACAATGCCAT	GCAGAGACAAAGAT 1656			
1b	1381	AGAGCAGAAAACGTAA	CAAGCAGCTTAGAGAG	AAAAAACAATGCCAT	GCAGAGACAAAGAT 1440			
1y	1657	AAAAAGGGGAAAAGGT	TAGAAACAGAAAAT	TAGAAAGGAAAATGGGT	TAATAAGAAAGAAAGCAC 1716			
1b	1441	AAAAAGGGGAAAAGGT	TAGAAACAGAAAAT	TAGAAAGGAAAATGGGT	TAATAAGAAAGAAAGCAC 1500			
1y	1717	AGAAATAAACTTCAGAC	AGCTGTCTTAAGAAA	ACAGGAGAGAGAGAGAG	AGGAACTTAAAGT 1776			
1b	1501	AGAAATAAACTTCAGAC	AGCTGTCTTAAGAAA	ACAGGAGAGAGAGAGAG	AGGAACTTAAAGT 1560			
1y	1777	GCAGCTAAAGAGAAAAG	CTCCAGAGACAAGCT	CTCTTCAAAAAAGAGAGAT	CAA 1836			
1b	1561	GCAGCTAAAGAGAAAAG	CTCCAGAGACAAGCT	CTCTTCAAAAAAGAGAGAT	CAA 1620			
1y	1837	AGATGAAAAAGATT	TAAAAAGGACAAAGAA	CCCCAAAGAGAGTTA	AGAGTTTATGGATCG 1896			
1b	1621	AGATGAAAAAGATT	TAAAAAGGACAAAGAA	CCCCAAAGAGAGTTA	AGAGTTTATGGATCG 1680			
1y	1897	AAAGAAAGGATT	TACAGAAAGTTAAG	TCGCAGAAATGAGAA	TTTCATGCCACCAAACTTAA 1956			
1b	1681	AAAGAAAGGATT	TACAGAAAGTTAAG	TCGCAGAAATGAGAA	TTTCATGCCACCAAACTTAA 1740			
1y	1957	ACATAC	TGAGAAATCTTT	CAGCGCCCTGGAGGAGGGCC	CAGCGTGCACACCAAGAGAGC 2016			
1b	1741	ACATAC	TGAGAAATCTTT	CAGCGCCCTGGAGGAGGGCC	CAGCGTGCACACCAAGAGAGC 1800			
1y	2017	TGAGGGCGCCCCCA	AGGTGGAAGCGGGCA	AAAGGCTGGAGAGCTT	CGTCGTCGTCGCGG 2076			
1b	1801	TGAGGGCGCCCCCA	AGGTGGAAGCGGGCA	AAAGGCTGGAGAGCTT	CGTCGTCGTCGCGG 1860			
1y	2077	GGAGACCGGAGAGCG	AGAGTTTCGAGAAAGCT	CTCAAAAGAGACAGCAG	GAGCGGCTTTGGA 2136			
1b	1861	GGAGACCGGAGAGCG	AGAGTTTCGAGAAAGCT	CTCAAAAGAGACAGCAG	GAGCGGCTTTGGA 1920			
1y	2137	GCTGGAGGAACT	CAAGAAAAGAGGAGG	AGAGAAAGGTCCT	GGAGAGGAGAAAGCA 2196			
1b	1921	GCTGGAGGAACT	CAAGAAAAGAGGAGG	AGAGAAAGGTCCT	GGAGAGGAGAAAGCA 1980			
1y	2197	GAGGAGGAGCAGG	AGGAGCCGATCGAA	AACTCAGAGAGGAGG	AGAGAGAGGAGGCT 2256			
1b	1981	GAGGAGGAGCAGG	AGGAGCCGATCGAA	AACTCAGAGAGGAGG	AGAGAGAGGAGGCT 2040			
1y	2257	AAAGGAGAGATT	TGAAAGGGCAAGAG	CAGAAAGCTGCT	GAGAAAACGCGCAGAGATGCCAGA 2316			

2041	AAAGGAAGAGATTGAAAGCGCGAAGAGCAGAAAGCTGCTGAGAAACGCCGCAAGAGTCCGAGA	2100
2317	AGATGGCTTTGTGTCAGATGACAAGAAACCAATTCAAGTGTTCCTCCTTAAAGGTTTCATCTCT	2376
2101	AGATGGCTTTGTGTCAGATGACAAGAAACCAATTCAAGTGTTCCTCCTTAAAGGTTTCATCTCT	2160
2377	CAAGATAGNAGACGCGACGAGAAATTTTGGATTAAGTCTGTGCGAGAAAGCAGTGGTGTCAA	2436
2161	CAAGATAGNAGACGCGACGAGAAATTTTGGATTAAGTCTGTGCGAGAAAGCAGTGGTGTCAA	2220
2437	ATCGACCCATCAAGCAGCAATATGCTCTCCAAGATTGACAGCAGACTCGAGCAGTATACCAG	2496
2221	ATCGACCCATCAAGCAGCAATATGCTCTCCAAGATTGACAGCAGACTCGAGCAGTATACCAG	2280
2497	TGCAATTGAGGGAAACAAAGAGCGCAAAACCTCAAGCCGGCGAGCCCTCGGATCTTCTCTGT	2556
2281	TGCAATTGAGGGAAACAAAGAGCGCAAAACCTCAAGCCGGCGAGCCCTCGGATCTTCTCTGT	2340
2557	TCTGCTGAAGTGTACGCAACATCAAGAGTATGTGCGGAGAAAGGGAATGTGTTTTCATC	2616
2341	TCTGCTGAAGTGTACGCAACATCAAGAGTATGTGCGGAGAAAGGGAATGTGTTTTCATC	2400
2617	CCCCACTGCAGCAGGCACACCCAAATGAAGAAACCTGCTGCGCTTGAAGGTAGGGGTTTCTAG	2676
2401	CCCCACTGCAGCAGGCACACCCAAATGAAGAAACCTGCTGCGCTTGAAGGTAGGGGTTTCTAG	2460
2677	CGCGATCAATGAATGGCTTAACATAAAACCCACAGATGCAAAACAAGTCACTGCTCCCAAAACC	2736
2461	CGCGATCAATGAATGGCTTAACATAAAACCCACAGATGCAAAACAAGTCACTGCTCCCAAAACC	2520
2737	TTCTGACTTTGAGACCCAGGAGACGTATCCAGCAAGCGGAACCTCTGGGAAAGCAATCTGT	2796
2521	TTCTGACTTTGAGACCCAGGAGACGTATCCAGCAAGCGGAACCTCTGGGAAAGCAATCTGT	2580
2797	GGATAGGTCACCTTCCCCCACTAAGGTTTGAGACAGTCCAGAAAGAAACCAAGCTCAAG	2856
2581	GGATAGGTCACCTTCCCCCACTAAGGTTTGAGACAGTCCAGAAAGAAACCAAGCTCAAG	2640
2857	ACGCAGCAGCAGCTCAGTTGTGAGAGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTT	2916
2641	ACGCAGCAGCAGCTCAGTTGTGAGAGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTT	2700
2917	ACTTAAATCGGTTTCATATCTCTTTTATTTTTCATATCCCAAGTAAACCCATGTATATTATC	2976
2701	ACTTAAATCGGTTTCATATCTCTTTTATTTTTCATATCCCAAGTAAACCCATGTATATTATC	2760
2977	ACTATATTTAAATACACAGCTCTAGAGATGTTTCATGTTAAAGTACTGCTTTCACAGG	3036
2761	ACTATATTTAAATACACAGCTCTAGAGATGTTTCATGTTAAAGTACTGCTTTCACAGG	2820
3037	AGCCTGTTTCTTAAAGAAAACCATGCTGTGAATATAGAGACTTTTCTACTGATCATCATAC	3096
2821	ATCCCTGTTTCTTAAAGAAAACCATGCTGTGAATATAGAGACTTTTCTACTGATCATCATAC	2880
3097	TCGTGATCTGAGCAGTGTATACCAACACATCTCGAAGTCAACAGAAAGATCCAAAGTTTAAAA	3156
2881	TCGTGATCTGAGCAGTGTATACCAACACATCTCGAAGTCAACAGAAAGATCCAAAGTTTAAAA	2940
3157	TTCCCTGCGGAATGTTGTCAGTATCTAGAAAAATGAACCGTAGTTTTTGTGTTTTTAAAT	3216
2941	TTG-CTGCGGAATGTTGTCAGTATCTAGAAAAATGAACCGTAGTTTTTGTGTTTTTAAAT	2999
3217	ACAGAGTCATGTTGTTCTGCACTTTTATAATAAAGCATGCAAGAAATTAATCTTAGTAGG	3276
3000	ACAGAGTCATGTTGTTCTGCACTTTTATAATAAAGCATGCAAGAAATTAATCTTAGTAGG	3059
3277	CAATTGTGAACACTTTTGTAAAGTAACCCATTTTGTGTTTTTGTGTTTTTAAAT	3336
3060	CAATTGTGAACACTTTTGTAAAGTAACCCATTTTGTGTTTTTGTGTTTTTAAAT	3119
3337	CTTTTAAAAAAAAGAAAAAGATCGGCGACAGTTATTCCTTT	3378
3120	TCCTTAAAAAAAAGAAATGTAAGTCTGTTAAGTATTAATCTTT	3161

XX 04-OCT-2001; 2001US-00974298.
XX 05-OCT-2000; 2000US-0238331P.
XX (CHEN/) CHEN H.
XX Chen H;
XX WPI; 2003-182653/18.
XX New cDNAs, which are differentially expressed in (metastatic) breast
XX cancer useful for diagnosing or staging, breast cancer, or for monitoring
XX the treatment of breast cancer in an individual.
XX Claim 1; SEQ ID NO 36; 30pp; English.
XX The invention describes a combination of cDNAs (designated BC-cDNAs),
XX which are differentially expressed in breast cancer. The combination
XX includes 152 cDNA sequences, or their complements. The protein encoded by
XX any of these BC-cDNAs is useful for screening several molecules or
XX compounds to identify at least one ligand that specifically binds the
XX protein, producing or preparing polyclonal or monoclonal antibodies, or
XX purifying antibodies from a sample. The antibodies, which specifically
XX bind the protein differentially expressed in breast cancer is useful for
XX detecting the expression of a protein in a sample. The BC-cDNAs are also
XX useful for diagnosing, monitoring the treatment of, or staging, breast
XX cancer. This sequence represents a differentially expressed breast cancer
XX associated cDNA. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from the US patent office at
XX seqdata.uspto.gov/sequence.html?DocID=20020156263
XX
XX Sequence 4547 BP; 1496 A; 849 C; 1033 G; 1169 T; 0 U; 0 Other;
XX
XX Query Match 48.4%; Score 1793.6; DB 7; Length 4547;
XX Best Local Similarity 76.9%; Pred. No. 0;
XX Matches 2585; Conservative 0; Mismatches 9; Indels 766; Gaps 2;
XX
QY 1 GCCTAGGGATGACAGGATCTCCACAGGAGGCTGCATCCACCTT-GGCTGGGGTGTGCG 59
DB 84 GCCTAGGGATGACAGGATCTCCACAGGAGGCTGCATCCACCTTGGGCTGGGGTGTGCG 143
QY 60 TCATTGGCTGCTTATTAGAAAACGACAGGACATGCATACACGGCTCCGACTGTAA 119
DB 144 TCATTGGCTGCTTATTAGAAAACGACAGGACATGCATACACGGCTCCGACTGTAA 203
QY 120 ACATAGGGGATATGTGTTCACTTAGCATGGACTTCTGGAGGGGCCAAGGAGGCGGTC 179
DB 204 ACATAGGGGATATGTGTTCACTTAGCATGGACTTCTGGAGGGGCCAAGGAGGCGGTC 263
QY 180 TGGAGTTTATTGAATAGAGAGGTGTATTCGGCTGCGCTGCCCGCTGCTTGTCTC 239
DB 264 TGGAGTTTATTGAATAGAGAGGTGTATTCGGCTGCGCTGCCCGCTGCTTGTCTC 323
QY 240 TCTGGCTGCTGCTGCTTAAAGAAATCAGTCTTCTTTCGACTTAGTCTCCTGGGAAG 299
DB 324 TCTGGCTGCTGCTGCTTAAAGAAATCAGTCTTCTTTCGACTTAGTCTCCTGGGAAG 383
QY 300 AAGTTTCAGACTACAAGGTATCATTTGGAACATTTCAAGATCATCAATCAAAATTCACAG 359
DB 384 AAGTTTCAGACTACAAGGTATCATTTGGAACATTTCAAGATCATCAATCAAAATTCACAG 443
QY 360 GGATTGTGTGACCAACAGAGGCTCAGACATCTGATTGCTGACCTGTCCAGACATCATCT 419
DB 444 GGATTGTGTGACCAACAGAGGCTCAGACATCTGATTGCTGACCTGTCCAGACATCATCT 503
QY 420 GGCTCCTGTAACCTGAATACACCATGTGATTTGAGCGTCCGACAGAACTTAGAA 479
DB 504 GGCTCCTGTAACCTGAATACACCATGTGATTTGAGCGTCCGACAGAACTTAGAA 563
XX 480 GGCAAAAGAGGGAGAGTGGACTCCGAGGAGAAAGAAATCGCTTACCAGAGGAATGACG 539

DB 564 GGC AAAAGAGGGAGGAGATGCGACTCTGAAGCAGAAAGAAATCGCTTACCAGAGGAATGACG 623
QY 540 ATGATGAAGAGGAGGACAGCCCGGGAAACGGCGCGCGAGCCCGACAGAGAAACGGCTGGCGC 599
DB 624 ATGATGAAGAGGAGGACAGCCCGGGAAACGGCGCGCGAGCCCGACAGAGAAACGGCTGGCGC 683
QY 600 AAGACGAGGAGGAGAAATCTTTGGGACACAGGTGACCGACCGAGGTGGAGGTGAATGCCCAGA 659
DB 684 AAGACGAGGAGGAGAAATCTTTGGGACAGGTGACCGACCGAGGTGGAGGTGAATGCCCAGA 743
QY 660 ACAGTGTCTCTGACGAGGAGGCAAGACCAACCACTCACTCACTCACTCACTCACTCACTCACT 719
DB 744 ACAGTGTCTCTGACGAGGAGGCAAGACCAACCACTCACTCACTCACTCACTCACTCACTCACT 803
QY 720 ATGAGGCGCATCTCTGAGCGCTTGGCTGGGTGAGGAAGACGCCAAAAACGCCCTTC 779
DB 804 ATGAGGCGCATCTCTGAGCGCTTGGCTGGGTGAGGAAGACGCCAAAAACGCCCTTC 863
QY 780 AGGAGGCTCTGAGCGGCGAGAGAGGAGTTTCGACCCCACTAATACAGATGCAAGTGTGCGC 839
DB 864 AGGAGGCTCTGAGCGGCGAGAGAGGAGTTTCGACCCCACTAATACAGATGCAAGTGTGCGC 923
QY 840 TCCCAAGCAGAGAGATGCAAAATGACACAGAGAAATGAAATACCGAGAAAGAGAGAA 899
DB 924 TCCCAAGCAGAGAGATGCAAAATGACACAGAGAAATGAAATACCGAGAAAGAGAGAA 983
QY 900 AAAGTGAAGTCCCAAGAAAGATACGAGATAGAGAAACAGAAACAGTCCCAAGTCT 959
DB 984 AAAGTGAAGTCCCAAGAAAGATACGAGATAGAGAAACAGAAACAGTCCCAAGTCT 1043
QY 960 ACCAGAGAAATGATTGGAGGGATGCTGAAGAAACAAAGAAAGAGAAAGAGAGAGG 1019
DB 1044 ACCAGAGAAATGATTGGAGGGATGCTGAAGAAACAAAGAAAGAGAAAGAGAGG 1103
QY 1020 AGGAGGAGAGAGAGAGCCAAAGCGAGGGAGCTTTGGAGAAATTCAGGTAGAGGTGATGG 1079
DB 1104 AGGAGGAGAGAGAGAGCCAAAGCGAGGGAGCTTTGGAGAAATTC 1148
QY 1080 TGGAGAGAAACAACTGAAAGCCAGGAGGAAACAGTGGTAACTCTATTAAAAAATGGC 1139
DB 1149 ----- 1148
QY 1140 AGATCAGTTTCAAGAGAGCTTAAACAGAGAGGAGAGGAAACAAAGTTTCAGATGAGATT 1199
DB 1149 ----- 1148
QY 1200 CCCATCATGAAAGATGGAAGAGGAAACAAAGAAAGAGCTGAGGACAGAGAGGCAAGT 1259
DB 1149 ----- 1148
QY 1260 TGGAGCAGAGAAAGAGAAAGAAATTAAGCCGAGCAAGACAAAGATAGCAGATGAAC 1319
DB 1149 ----- 1148
QY 1320 GAGCAAGATTGAAGCAGAGAAAGAAAGCAGCTGCCAAGAAAGAGAAAGAGAGAGGAGCAG 1379
DB 1149 ----- 1148
QY 1380 AAGAGAGGAAAGGATGAGGGAGGAGAAAGAAAGGCGCAGAGAGAGAGGAGAGATAA 1439
DB 1149 ----- 1148
QY 1440 AGGAGAGAGAAAGGCGCAGCAGAGAGGAGGAGGATTAAGGAGGAGAGAAAGAGG 1499
DB 1149 ----- 1148
QY 1500 CAGCAGAGGAGAGGCGCAGAGGATAAAGAGGAGAGAAAGGCGCAGCAGAGAGAGAGCAAA 1559
DB 1149 ----- 1148
QY 1560 GGGCCGCGGAGGAGGAGAGAGAGAGGCTAAGGTAGAGAGGAGAGAAACCTTAACAGCAGC 1619
DB 1149 ----- 1148

PF	25-MAR-1992;	92EP-00302591.	QY	1961	ACTGAGATACCTTTTCAGCGCCCTCGAGGAGGCGCCAGCGTGGACACACAGGAGGCTGAG	2020
XX	29-MAR-1991;	91JP-00089106.	DB	822	ACTGAGATACCTTTTCAGCGCCCTCGAGGAGGCGCCAGCGTGGACACACAGGAGGCTGAG	881
PR	27-DEC-1991;	91JP-00358040.	QY	2021	GGCGCCCCCAGGTGGAAGCCCGGCAAAAGGCTGGAGGAGCTTCGTCTCTCGCGGGAG	2080
XX	(SOBU/)	SObue K.	DB	882	GGCGCCCCCAGGTGGAAGCCCGGCAAAAGGCTGGAGGAGCTTCGTCTCTCGCGGGAG	941
PA	(TAKI)	TAKARA SHUZO CO LTD.	QY	2081	ACCGAGACGGAAGAGTTTGGAGAGCTTCAACACAGACAGCAGGAGGCGGCTTTGGAGCTG	2140
XX	Hayashi K, Asada K, Hashida T, Kotani H, Kato I, Sobue K;		DB	942	ACCGAGACGGAAGAGTTTGGAGAGCTTCAACACAGACAGCAGGAGGCGGCTTTGGAGCTG	1001
XX	WPI; 1992-325651/40.		QY	2141	CAGGAACCTCAAGAAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2200
XX	Polypeptide(s) with calmodulin- and actin binding activities - useful for		DB	1002	CAGGAACCTCAAGAAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1061
XX	treating and diagnosing cancer and regulating digestive tract motility as		QY	2201	AGGAAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2260
XX	vasodilators.		DB	1062	AGGAAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1121
XX	Disclosure; Page 33-34; 47pp; English.		QY	2261	GAAGAGATTGAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2320
XX	cDNA clones that encode polypeptide of human caldesmon were selected from		DB	1122	GAAGAGATTGAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1181
XX	a cDNA library prep'd. from HeLa cells, and the DNA sequence and the amino		QY	2321	GGCTTGTGATGACAAAGAACCATTCAGAGTGTTCACCTCAAGAGTTCATCTCTCAAG	2380
XX	acid sequence encoding human caldesmon were deduced by DNA sequence		DB	1182	GGCTTGTGATGACAAAGAACCATTCAGAGTGTTCACCTCAAGAGTTCATCTCTCAAG	1241
XX	analysis. Next, the inventors prepared polypeptides of various lengths,		QY	2381	ATAGAAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2440
XX	and looked for the functional domain of human caldesmon polypeptide,		DB	1242	ATAGAAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1301
XX	using calmodulin-binding activity and actin-binding activity as indices.		QY	2441	ACCCATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2500
XX	There are two forms of caldesmon. The form of higher mol. wt. is called		DB	1302	ACCCATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1361
XX	type I and the form of lower mol. wt. is called type II. 312AA is a		QY	2501	ATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2560
XX	polypeptide that is the C-terminal portion of types I and II. 122AA and		DB	1362	ATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1421
XX	118A are polypeptides that have a sequence of 115 AAs in the C-terminal		QY	2561	GCTGAAGGTGTACGCAACATCAAGAGTGTGTGGAGGAGGAGGAGGAGGAGGAGGAGG	2620
XX	sequence of 312AA, and 94AA and 90AA are polypeptides that have a		DB	1422	GCTGAAGGTGTACGCAACATCAAGAGTGTGTGGAGGAGGAGGAGGAGGAGGAGGAGG	1481
XX	sequence of 88 AAs that are in the central portion of 118AA, lacking		QY	2621	ACTGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2680
XX	both the N-terminal portion and the C-terminal portion. On the N-terminal		DB	1482	ACTGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1541
XX	of 118AA, there are residues Met-Ala that originated from a PCR primer,		QY	2681	ATCAATGAATGGCTAACTAAACCCAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2740
XX	and on the N-terminal of 90AA, there are residues Met-Ala that also		DB	1542	ATCAATGAATGGCTAACTAAACCCAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1601
XX	originated from a PCR primer. On the N-terminal of 122AA and 94AA, there		QY	2741	GACTGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2800
XX	are residues Met-Ala that originated from PCR primer. The polypeptide		DB	1602	GACTGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1661
XX	(typeI, 312AA, 122AA, and 116AA) that contains the		QY	2801	AAGGTCACTTCCCGCCACTAAGGTTTGAGACAGTTTCCAGAAAGAACCCCAAGCTCAAGAG	2860
XX	polypeptide of 116 AAs has both calmodulin-binding activity and actin-		DB	1662	AAGGTCACTTCCCGCCACTAAGGTTTGAGACAGTTTCCAGAAAGAACCCCAAGCTCAAGAG	1721
XX	binding activity. Type I, type II, 312AA, 122AA, 118AA, 116AA all have		QY	2861	AGGACGAGCTCAGTTGTAGAGGCTAAATTCGCTCTGTTTGTATTTATTTGTTGATTTACTA	2920
XX	tropomyosin-binding activity and have an inhibitory activity to		DB	1722	AGGACGAGCTCAGTTGTAGAGGCTAAATTCGCTCTGTTTGTATTTATTTGTTGATTTACTA	1779
XX	actomyosin ATPase, and so the functional unit of human caldesmon is		QY	2921	AATTGGGTTTCATTATCTTTTATTTTCAATATCCAGTAAACCCATGTATATATCACTA	2980
XX	identified as being 116AA. (Updated on 25-MAR-2003 to correct FN field.)		DB	1780	AATTGGGTTTCATTATCTTTTATTTTCAATATCCAGTAAACCCATGTATATATCACTA	1839
XX	(Updated on 25-MAR-2003 to correct PA field.)		QY	2981	TATTTAATATCAAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTTGCACAGAGCC	3040
XX	Sequence 2215 BP; 778 A; 436 C; 593 G; 407 T; 0 U; 1 Other;		DB	1840	TATTTAATATCAAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTTGCACAGAGCC	1898
XX	Query Match 42.3%; Score 1567; DB 2; Length 2215;		QY	3041	TGTTTTCTAAAGAAACCCATCTGCTGTGAAATAGAGACTTTTCTACTGATCATATACTCTG	3100
XX	Best Local Similarity 94.4%; Pred. No. 0;					
XX	Matches 1658; Conservative 0; Mismatches 95; Indels 4; Gaps 3;					
QY	1601	CAGAACGTAACAGAGCAGTGTAGAGAGAAACATGCCATGCAAGAGCAAAAGATATAAA	1660			
DB	462	CAAGAAAGATACAGATAGAGGAAACAGAACAGTCCAAAGTCTTACCAGAGATGAT	521			
QY	1661	GGGGAAGAGGTAGAACAGAGAAATAGAGGGAAATGGGTAAATGAAAGAAAGCAAGAA	1720			
DB	522	TGGAGGATCTGAGAGAAACAGAAAGAGACAGCAAGGAAAGGAGGAGGAGGAGGAG	581			
QY	1721	GATAAAGTTCAGACAGTGTCTTAAGAACAGGAGGAGGAGGAGGAGGAGGAGGAGG	1780			
DB	582	AAGCCAAAGCGAGGAGCATTTGGAGAAATCAGGAGAGAGGAGGAGGAGGAGGAGG	641			
QY	1781	GCTTAAAGAGAGAAAGCTCCCAAGAGAGCAAGGCTTACCTTCAAAAAAGAGAGATCA	1840			
DB	642	GCTTAAAGAGAGAAAGCTCCCAAGAGAGCAAGGCTTACCTTCAAAAAAGAGAGATCA	701			
QY	1841	GAAAGATTAAAGGAGCAAGAACCCCAAGAGAGTAAAGTTCATGGATCGAAG	1900			
DB	702	GAAAGATTAAAGGAGCAAGAACCCCAAGAGAGTAAAGTTCATGGATCGAAG	761			
QY	1901	AAGGATTTACAGAGTTAAGTTCAGAGATCGAGATTCATGACCCCAAACTTAAACAT	1960			
DB	762	AAGGATTTACAGAGTTAAGTTCAGAGATCGAGATTCATGACCCCAAACTTAAACAT	821			

2292	CTGAGAAACGCCAGAGATGCCAGAGATGGCTTGTTCAGATGACAAGAAACCATTCAGT	2351
1075	CTGAGAAACGCCAGAGATGCCAGAGATGGCTTGTTCAGATGACAAGAAACCATTCAGT	1134
2352	GTTTCACCTCTAAAGGTTTCATCTCTCAAGATAGAAGAGCGAGCAGAAATTTTGTGAATAAGT	2411
1135	GTTTCACCTCTAAAGGTTTCATCTCTCAAGATAGAAGAGCGAGCAGAAATTTTGTGAATAAGT	1194
2412	CTGTGCAGAAAGCGAGTGGTGTCAAAATCGACCCATCAAGCGAGCAATAGTCTCCAAGTTG	2471
1195	CTGTGCAGAAAGCGAGTGGTGTCAAAATCGACCCATCAAGCGAGCAATAGTCTCCAAGTTG	1254
2472	ACAGCAGACTGGAGCAGTATACAGATGCAATTGAGGGAAACAAAAGCCAAAACCTACAA	2531
1255	ACAGCAGACTGGAGCAGTATACAGATGCAATTGAGGGAAACAAAAGCCAAAACCTACAA	1314
2532	AGCCGGCAGCTCGGATCTCTCTGTTCTCTGTTGAGGTGTACGACACATCAAGATATGT	2591
1315	AGCCGGCAGCTCGGATCTCTCTGTTCTCTGTTGAGGTGTACGACACATCAAGATATGT	1374
2592	GGGAGAAAGGGAATGTGTTTTCATCCCCCATCTGCAGCAGGCACACCAATTAAGGAAACTG	2651
1375	GGGAGAAAGGGAATGTGTTTTCATCCCCCATCTGCAGCAGGCACACCAATTAAGGAAACTG	1434
2652	CTGGCTTGAAGTAGGGTTTCTAGCCGCATCAATGAATGCTTAACAAAACCCAGATG	2711
1435	CTGGCTTGAAGTAGGGTTTCTAGCCGCATCAATGAATGCTTAACAAAACCCAGATG	1494
2712	GAAAACAAGTCACCTGTCTCCCAAACCTTCTGACTTGAGACGAGGACGATATCCAGCAAGC	2771
1495	GAAAACAAGTCACCTGTCTCCCAAACCTTCTGACTTGAGACGAGGACGATATCCAGCAAGC	1554
2772	GGAACTCTGGGHAAGACATCTGTGATAGGTCATCTCCCCACTAAGGTTTGAGACA	2831
1555	GGAACTCTGGGHAAGACATCTGTGATAGGTCATCTCCCCACTAAGGTTTGAGACA	1614
2832	GTTTCAGAAAGAACCCAAAGCTCAAGAGCGAGGACGAGCTCAGTTGTAGAGGCTAATTCG	2891
1615	GTTTCAGAAAGAACCCAAAGCTCAAGAGCGAGGACGAGCTCAGTTGTAGAGGCTAATTCG	1674
2892	CTCTGTTTTGTATTATGTTGATTTATCAATTCGTTTTCATTTATTTTTCATAA	2951
1675	CTC--TTTTGTATTATGTTGATTTATCAATTCGTTTTCATTTATTTTTCATAA	1732
2952	TCCCAGTAAACCCATGATATTTATCATATATTTAATAATCACAGCTAGAGATGTCAT	3011
1733	TCCCAGTAAACCCATGATATTTATCATATATTTAATAATCACAG-CTAGAGATGTCAT	1791
3012	GGTAAAGTACTGCCCTTGCACGAGGCTCTTCTTAAGAAACCCATCCTGTGGAATAG	3071
1792	GGTAAAGTACTGCCCTTGCACGAGGCTCTTCTTAAGAAACCCATCCTGTGGAATAG	1851
3072	AGACTTTTCTACTGATCATATACTCTGTATCTGCAGCGTGAATACCAACCATCTGAA	3131
1852	AGACTTTTCTACTGATCATATAACTCTGAATCTGCAGCGTGAATACCAACCATCTGAA	1911
3132	GTCAACAGAGATCCAGTTTAAATTCGCTCGGAATGTGTCAGTATCTAGAAAATG	3191
1912	GTCAACAGAGATCCAGTTTAAATTCGCTCGGAATGTGTCAGTATCTAGAAAATG	1971
3192	AACCGTAGTTTTGTTTTTAAATACAGAAGTCATGTTGTTTCTGCACTTTATATAAA	3251
1972	AACCGTAG--TTTTGTTTTTAAATACAGAAGTCATGTTGTTTCTGCACTTTATAATAA	2030
3252	GCATGGAAGAAATCTTACTAGGCAATGTAAACACTTTTTTGAAGTAGTAAACCATTCAG	3311
2031	GCATGGAAGAAATCTTACTAGGCAATGTAAACACTTTTTTGAAGTAGTAAACCATTCAG	2090
3312	ATTGGAATACTCGGATAATGGTCTTTAAAAAAAAGAAAAAGA	3357
2091	ATTGGAATACTCGAATAATGGTCTTTAAAAAAAAGAAAAAAA	2136

RESULT 8	
AAQ28986	
ID	AAQ28986 standard; cDNA to mRNA; 1677 BP.
XX	
AC	AAQ28986;
XX	
XX	
DT	25-MAR-2003 (revised)
DT	25-FEB-1993 (first entry)
XX	
DE	Sequence encoding a polypeptide of human caldesmon designated type I
DE	(heavy) AA.
XX	
KW	Caldesmon; calmodulin; actin; binding; actomysin system; regulation; ss.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FT	1..1677
CDS	/*tag= a
FT	
XX	
PN	EP506377-A2.
XX	
PD	30-SEP-1992.
XX	
PF	25-MAR-1992; 92EP-00302591.
XX	
PR	29-MAR-1991; 91JP-00089106.
PR	27-DEC-1991; 91JP-00358040.
XX	
PA	(SOBU/) SOBUE K.
PA	(TAKI) TAKARA SHUZO CO LTD.
XX	
PI	Hayashi K, Asada K, Hashida T, Kotani H, Kato I, Sobue K;
XX	
DR	WPI; 1992-325651/40.
DR	P-PSDB; AAR27363.
XX	
PT	Polypeptide(s) with calmodulin- and actin binding activities - useful for
PT	treating and diagnosing cancer and regulating digestive tract motility as
PT	vasodilators.
XX	
PS	Claim 5: Page 29-30: 47pp; English.

cDNA clones that encode polypeptide of human caldesmon were selected from a cDNA library prep'd. from HeLa cells, and the DNA sequence and the amino acid sequence encoding human caldesmon were deduced by DNA sequence analysis. Next, the inventors prepared polypeptides of various lengths, and looked for the functional domain of human caldesmon polypeptide, using calmodulin-binding activity and actin-binding activity as indices. There are two forms of caldesmon. The form of higher mol. wt. is called type I and the form of lower mol. wt. is called type II. 312AA is a polypeptide that is the C-terminal portion of types I and II. 122AA and 118A are polypeptides that have a sequence of 116 AAs in the C-terminal sequence of 312AA, and 94AA and 90AA are polypeptides that have a sequence of 88 AAs that are in the central portion of 118AA, lacking both the N-terminal portion and the C-terminal portion. On the N-terminal of 118AA, there are residues Met-Ala that originated from a PCR primer, and on the N-terminal of 90AA, there are residues Met-Ala that also originated from a PCR primer. On the N-terminal of 122AA and 94AA, there are residues Met-Ala that originated from PCR primer. The polypeptide (typeI, typeII, 312AA, 122AA, 118AA, and 116AA) that contains the polypeptide of 116 AAs has both calmodulin-binding activity and actin-binding activity. Type I, type II, 312AA, 122AA, 118AA, 116AA all have tropomyosin-binding activity and have an inhibitory activity to actomyosin ATPase, and so the functional unit of human caldesmon is identified as being 116AA. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to correct PA field.)

XX	Sequence 1677 BP; 594 A; 345 C; 501 G; 237 T; 0 U; 0 Other;
SQ	
Query Match	29.3%; Score 1086.2; DB 2; Length 1677;
Best Local Similarity	92.8%; Pred. No. 7.6e-244;

QY	2681	ATCAATGAATGGCTAACTAAAAACCCAGATGGAAACAAGTCACCTGCTCCCAACCTTCT	2740
Db	1531	ATCAATGAATGGCTAACTAAAAACCCAGATGGAAACAAGTCACCTGCTCCCAACCTTCT	1590
QY	2741	GACTTGAGACCGAGACGATATCCAGCAACGCGAACCCTCTGGGAAAGCAATCTGTGGAT	2800
Db	1591	GACTTGAGACCGAGACGATATCCAGCAACGCGAACCCTCTGGGAAAGCAATCTGTGGAT	1650
QY	2801	AAGGTCACCTTCCCCCCTCACTAAGGTTTGA	2827
Db	1651	AAGGTCACCTTCCCCCCTCACTAAGGTTTGA	1677
RESULT 9			
AAQ28985			
ID ID AAQ28985 standard; cDNA to mRNA; 1599 BP.			
XX	AC	AAQ28985;	
XX	AC	AC	
XX	DT	25-MAR-2003 (revised)	
DT	DT	25-FEB-1993 (first entry)	
XX	XX	Sequence encoding a polypeptide of human caldesmon designated type II	
DE	DE	(light) AA.	
XX	XX	Caldesmon; calmodulin; actin; binding; actomyosin system; regulation; ss.	
XX	XX	Homo sapiens.	
XX	XX	Key	Location/Qualifiers
FT	FT	CDS	1..1599
FT	FT		/*tag= a
XX	XX	EP506377-A2.	
XX	XX	30-SEP-1992.	
PD	PD	25-MAR-1992; 92EP-00302591.	
XX	XX	29-MAR-1991; 91JP-00089106.	
PR	PR	27-DEC-1991; 91JP-00358040.	
XX	XX	(SOBU/) SOBUE K.	
PA	PA	(TAKI) TAKARA SHUZO CO LTD.	
XX	XX	Hayashi K, Asada K, Hashida T, Kotani H, Kato I, Sobue K;	
XX	XX	WPI; 1992-325651/40.	
DR	DR	P-PSDB; AAR27362.	
XX	XX	Polypeptide(s) with calmodulin- and actin binding activities - useful for	
PT	PT	treating and diagnosing cancer and regulating digestive tract motility as	
PT	PT	vasodilators.	
XX	XX	Claim 5; Page 27-28; 47pp; English.	
PS	PS	cDNA clones that encode polypeptide of human caldesmon were selected from	
XX	XX	a cDNA library prepd. from HeLa cells, and the DNA sequence and the amino	
CC	CC	acid sequence encoding human caldesmon were deduced by DNA sequence	
CC	CC	analysis. Next, the inventors prepared polypeptides of various lengths,	
CC	CC	and looked for the functional domain of human caldesmon polypeptide,	
CC	CC	using calmodulin-binding activity and actin-binding activity as indices.	
CC	CC	There are two forms of caldesmon. The form of higher mol. wt. is called	
CC	CC	type I and the form of lower mol. wt. is called type II. 312AA is a	
CC	CC	polypeptide that is the C-terminal portion of types I and II. 122AA and	
CC	CC	118A are polypeptides that have a sequence of 116 AAs in the C-terminal	
CC	CC	sequence of 312AA, and 94AA and 90AA are polypeptides that have a	
CC	CC	sequence of 88 AAs that are in the central portion of 118AA, lacking	
CC	CC	both the N-terminal portion and the C-terminal portion. On the N-terminal	
CC	CC	of 118AA, there are residues Met-Ala that originated from a PCR primer,	
CC	CC	and on the N-terminal of 90AA, there are residues Met-Ala that also	
CC	CC	originated from a PCR primer. On the N-terminal of 122AA and 94AA, there	

1364 GGGAGAAAGGGAATGTGTTTTCATCCCCCACTGCAGCAGGCACACCAAAATAGGAAACATG 1421

2652 CTGGCTTGAAGGTAGGGGTTTCTAGCCGCATCAATGAATGGCTAACTAAAACCCCGAGATG 2711

1424 CTGGCTTGAAGGTAGGGGTTTCTAGCCGATCAATGAATGGCTAACTAAAACCCCGAGATG 1483

2712 GAAACAAGTCACTGCTCCCAACCTTTCAGCTTGAGACACGAGAGCGTATCCAGCAAGC 2771

1484 GAAACAAGTCACTGCTCCCAACCTTTCAGCTTGAGACACGAGAGCGTATCCAGCAAGC 1543

2772 GGAACTCTGGGAAAAGCAATCTGTGATAAGGTCACTTCCCCCCTAAAGGTTTGA 2827

1544 GGAACTCTGGGAAAAGCAATCTGTGATAAGGTCACTTCCCCCCTAAAGGTTTGA 1599

RESULT 10
AAQ28984
ID AAQ28984 standard; cDNA to mRNA; 939 BP.

XX AC AAQ28984;

XX AC

XX DT 25-MAR-2003 (revised)

XX DT 25-FEB-1993 (first entry)

XX DE Sequence encoding a polypeptide of human caldesmon designated 312 AA.

XX KW Caldesmon; calmodulin; actin; binding; actomyosin system; regulation; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..939

XX FT /*tag= a

XX FT

XX PN EP506377-A2.

XX PD 30-SEP-1992.

XX PF 25-MAR-1992; 92EP-00302591.

XX PR 29-MAR-1991; 91JP-00089106.

XX PR 27-DEC-1991; 91JP-00358040.

XX PR

XX PA (SOBU/) SOBUE K.

XX PA (TAKI) TAKARA SHUZO CO LTD.

XX PI Hayaashi K, Asada K, Hashida T, Kotani H, Kato I, Sobue K;

XX PI

XX DR WPI; 1992-325651/40.

XX DR P-PSDB; AAR27361.

XX PT Polypeptide(s) with calmodulin- and actin binding activities - useful for treating and diagnosing cancer and regulating digestive tract motility as vasodilators.

XX PT

XX PS Claim 5; Page 26; 47pp; English.

XX PS

XX CC cDNA clones that encode polypeptide of human caldesmon were selected from a cDNA library prepd. from HeLa cells, and the DNA sequence and the amino acid sequence encoding human caldesmon were deduced by DNA sequence analysis. Next, the inventors prepared polypeptides of various lengths, and looked for the functional domain of human caldesmon polypeptide, using calmodulin-binding activity and actin-binding activity as indices. There are two forms of caldesmon. The form of higher mol. wt. is called type I and the form of lower mol. wt. is called type II. 312AA is a polypeptide that is the C-terminal portion of types I and II. 122AA and 118AA are polypeptides that have a sequence of 116 AAs in the C-terminal sequence of 312AA, and 94AA and 90AA are polypeptides that have a sequence of 88 AAs that are in the central portion of 118AA, lacking both the N-terminal portion and the C-terminal portion. On the N-terminal of 118AA, there are residues Met-Ala that originated from a PCR primer, and on the N-terminal of 90AA, there are residues Met-Ala that also originated from a PCR primer. On the N-terminal of 122AA and 94AA, there

CC are residues Met-Ala that originated from PCR primer. The polypeptide
CC (typeI, typeII, 312AA, 122AA, 118AA, and 116AA) that contains the
CC polypeptide of 116 AAs has both calmodulin-binding activity and actin-
CC binding activity. Type I, type II, 312AA, 122AA, 118AA, 116AA all have
CC tropomyosin-binding activity and have an inhibitory activity to
CC actomyosin ATPase, and so the functional unit of human caldesmon is
CC identified as being 116AA. (Updated on 25-MAR-2003 to correct PN field.)
XX (Updated on 25-MAR-2003 to correct PA field.)
SQ Sequence 939 BP; 314 A; 194 C; 275 G; 156 T; 0 U; 0 Other;

Query Match 25.2%; Score 932.6; DB 2; Length 939;
Best Local Similarity 99.6%; Pred. No. 5.5e-208;
Matches 935; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1889 ATGGATCGAAGAGGAGTTTACAGAAGTTAAGTCGCAGAAATGAGAAATCATGACCCAC 1948
Db 1 ATGGATCGAAGAGGAGTTTACAGAAGTTAAGTCGCAGAAATGAGAAATCATGACCCAC 60

QY 1949 AAATTTAAACATACTGAGAAATCTTTTCAGCGCCCTGGAGGAGCCAGGTCGACACC 2008
Db 61 AAATTTAAACATACTGAGAAATCTTTTCAGCGCCCTGGAGGAGCCAGGTCGACACC 120

QY 2009 AAGGAGGCTGAGGCGCCGCCAGGTCGAAAGCCGCAAAAGGCTGGAGGAGCTTCGTCGT 2068
Db 121 AAGGAGGCTGAGGCGCCGCCAGGTCGAAAGCCGCAAAAGGCTGGAGGAGCTTCGTCGT 180

QY 2069 CFTGCGGGGAGACCGAGAGCGAAGAGTTTCAGAAAGCTCAACAGAGCAGCAGGAGCG 2128
Db 181 CFTGCGGGGAGACCGAGAGCGAAGAGTTTCAGAAAGCTCAACAGAGCAGCAGGAGCG 240

QY 2129 GCTTTGAGGCTGAGGAACTCAAGAAAAGAGGAGGAGAGAGAGTCTCTGGAGGAG 2188
Db 241 GCTTTGAGGCTGAGGAACTCAAGAAAAGAGGAGGAGAGAGTCTCTGGAGGAG 300

QY 2189 GAAGACGAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2248
Db 301 GAAGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360

QY 2249 AGAGGCTAAGGAGAGAGATTGAAAGCGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAG 2308
Db 361 AGAGGCTAAGGAGAGAGATTGAAAGCGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAG 420

QY 2309 ATGCCAGAGAGATGCTTGTTCAGATGACAGAAACCAATTCAGATGTTTCACTCTAAGGT 2368
Db 421 ATGCCAGAGAGATGCTTGTTCAGATGACAGAAACCAATTCAGATGTTTCACTCTAAGGT 480

QY 2369 TCATCTCTCAGATAGAGAGCGAGCAATTTTGAATAAGTCTGTGAGAGAAAGCAGT 2428
Db 481 TCATCTCTCAGATAGAGAGCGAGCAATTTTGAATAAGTCTGTGAGAGAAAGCAGT 540

QY 2429 GGTGTCAAATCGACCCATCAAGCAGCAATAGTCTCAAGATTGACAGAGACTGGAGCAG 2488
Db 541 GGTGTCAAATCGACCCATCAAGCAGCAATAGTCTCAAGATTGACAGAGACTGGAGCAG 600

QY 2489 TATACCAAGTCAATGAGGAGCAAAAGCGCAAACTTACAAAGCCGACGCTCGGAT 2548
Db 601 TATACCAAGTCAATGAGGAGCAAAAGCGCAAACTTACAAAGCCGACGCTCGGAT 660

QY 2549 CTTCTCTGTTCTGCTGAAGGTTGACGCAACATCAAGAGTATGTGGAGAGAAAGGATGTG 2608
Db 661 CTTCTCTGTTCTGCTGAAGGTTGACGCAACATCAAGAGTATGTGGAGAGAAAGGATGTG 720

QY 2609 TTTTCATCCCCCATGTCAGAGCGCACCAATATAGGAATGCTGCTGTTGAGGTAGGG 2668
Db 721 TTTTCATCCCCCATGTCAGAGCGCACCAATATAGGAATGCTGCTGTTGAGGTAGGG 780

QY 2669 GTTTCAGCGCATCAATGAATGGTGTAACTAAACCCCGAGATGGAACAAAGTCACTGTG 2728
Db 781 GTTTCAGCGCATCAATGAATGGTGTAACTAAACCCCGAGATGGAACAAAGTCACTGTG 840

QY 2729 CCCAAACCTTCTGATTTGAGACAGAGAGCGTATCCAGAGGCGGAGACCTCTCGGAAAG 2788

Db 841 CCCAAACCTTCTGACTTGAGACCAGGAGAGCTATCCAGCAAGCGGACCTCTGGGAAAG 900
QY 2789 CAATCTGTGGATAAGTCACTTCCCCCACTAAGGTTTGA 2827
Db 901 CAATCTGTGGATAAGTCACTTCCCCCACTAAGGTTTGA 939

RESULT 11
AAC00905
ID AAC00905 standard; cDNA; 705 BP.
XX AAC00905;
AC AAC00905;
XX 06-OCT-2000 (first entry)
DT 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 903.
DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
OS Homo sapiens.
PN EP1033401-A2.
XX 06-SEP-2000.
PD 06-SEP-2000.
XX 21-FEB-2000; 2000BP-00200610.
PF 21-FEB-2000; 2000BP-00200610.
PR 26-FEB-1999; 99US-0122487P.
XX 26-FEB-1999; 99US-0122487P.
PA (GEST) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI WPI; 2000-500381/45.
XX P-PSDB; AAG00899.
DR New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
PS Claim 1; SEQ ID NO 903; 71pp + Sequence Listing; English.
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
XX vectors
SQ Sequence 705 BP; 190 A; 171 C; 200 G; 143 T; 0 U; 1 Other;

Query Match 18.7%; Score 691; DB 3; Length 705;
Best Local Similarity 99.9%; Pred. No. 1.8e-151;
Matches 691; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCTAGGAATGACAGGAGCTTCCACAGGAGGCTGCATCCACTTGGTGGGCTGTCGT 60
Db 14 GCCTAGGAATGACAGGAGCTTCCACAGGAGGCTGCATCCACTTGGTGGGCTGTCGT 73

QY 61 CATTGGCTGCTATTAGAAAAACGACAGAGCAATGCATACCAACCGCTCCCGACTGTAAA 120
Db 74 CATTGGCTGCTATTAGAAAAACGACAGAGCAATGCATACCAACCGCTCCCGACTGTAAA 133

QY 121 CATAGGGATATGTGTTTCATCTAGCATGACTTCTGGGAGGGGCCAAGGAAGCGGCTCT 180

CX 21-SEP-2000.
 CX 08-MAR-2000; 2000WO-US005882.
 CX 12-MAR-1999; 99US-0124270P.
 CX (HUMA-) HUMAN GENOME SCI INC.
 CX Rosen CA, Ruben SM;
 CX WPI: 2000-587533/55.
 CX P-PSDB; AAB43497.
 CX Novel isolated nucleic acids comprising sequences encoding peptides
 CX useful for treating or diagnosing e.g. cancer.
 CX Claim 1; Page 690-691; 2352pp; English.
 CX AAC77607 to AAC78448 encode the human cancer associated proteins given in
 CX AAB43398 to AAB44239. The proteins can have activities based on the
 CX tissues and cells the genes are expressed in. Example of activities
 CX include: cytostatic; proliferative; vulnery; immunomodulator;
 CX antidiabetic; antiaschemic; antirheumatic; antithrombotic; antitumor;
 CX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CX dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CX neotropic; vasotropic; antipsoriatic and antiangiogenic. The
 CX polynucleotides and polypeptides can be used for preventing, treating or
 CX ameliorating medical conditions and diagnosing pathological conditions.
 CX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CX or inhibiting the proliferation, differentiation or mobilisation of
 CX immune cells, to treat disorders of haematopoietic cells, autoimmune
 CX disorders, allergic reactions, graft versus host disease and organ
 CX rejection, modulate haemostatic or thrombolytic activity, modulate
 CX inflammation, cancers, cardiovascular disorders, neurological disease and
 CX bacterial or viral infections. The peptides, nucleotides, antibodies,
 CX agonists and antagonists may be also be used in drug screens. AAC78449 to
 CX AAC78457 and AAB44240 represent sequences used in the exemplification of
 CX the present invention
 CX Sequence 2511 BP; 814 A; 445 C; 443 G; 803 T; 0 U; 6 Other;
 CX
 Query Match 15.8%; Score 584.2; DB 3; Length 2511;
 Best Local Similarity 98.7%; Pred. No. 3.3e-126;
 Matches 589; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 2763 CCAGCAAGCGGAACCTCTGGGAAAGCAATCTGTGGATAAGTCACTTCCCCCACTAAGG 2822
 73 CGAGCAAGCGGAACCTCTGGGAAAGCAATCTGTGGATAAGTCACTTCCCCCACTAAGG 132
 2823 TTTGAGCAGTTCAGAAAGAACCCAGCTCAAGCAGCAGCAGTCACTTGTAGAGG 2882
 133 TTTGAGCAGTTCAGAAAGAACCCAGCTCAAGCAGCAGCAGTCACTTGTAGAGG 192
 2883 GCTAATTCGCTCTGTTTGTATTTATGTTGATTAATTAATTAATTAATTAATTAATTAAT 2942
 193 GCTAATTCGCTCTGTTTGTATTTATGTTGATTAATTAATTAATTAATTAATTAATTAAT 252
 2943 TTTTCAATATCCAGTAACCCAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3002
 253 TTTTCAATATCCAGTAACCCAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 312
 3003 GATGTTTCATGTAAGTACTGCTTTTGCAGGAGCCGTTTCTTAAGAAACCCATGCT 3062
 313 GATGTTTCATGTAAGTACTGCTTTTGCAGGAGCCGTTTCTTAAGAAACCCATGCT 372
 3063 GTGAATATAGAGACTTTTCTACTGATCATCAATCACTCTGATCTGAGCAGTGTATCAAC 3122
 373 GTGAATATAGAGACTTTTCTACTGATCATCAATCACTCTGATCTGAGCAGTGTATCAAC 432
 3123 ACATCTGAAGTCAACAGAGATCCAAAGTTTAAATTTGCTCGGAATGTGTGAGTATCT 3182

Db 433 ACATCTGAAGTCAACAGAGATCCAAAGTTTAAATTTGCTCGGAATGTGTGAGTATCT 492
 Qy 3183 AGAAAAATGAACCGTAGTTTGTGTTTAAATTAACAGAGTCACTGTGTTTCTGCACCT 3242
 Db 493 AGAAAAATGAACCGTAGTTTGTGTTTAAATTAACAGAGTCACTGTGTTTCTGCACCT 552
 Qy 3243 TATAATAAGCATGGAAGAAATTAATCTTAGTAGGCAATTTAATCACTTTTGAAGTAAC 3302
 Db 553 TATAATAAGCATGGAAGAAATTAATCTTAGTAGGCAATTTAATCACTTTTGAAGTAAC 612
 Qy 3303 CCATTTGAGATTTGAAATTAATCTTAGTAGGCAATTTAATCACTTTTGAAGTAAC 3359
 Db 613 CCATTTGAGATTTGAAATTAATCTTAGTAGGCAATTTAATCACTTTTGAAGTAAC 669
 RESULT 15
 AAH33117
 ID AAH33117 standard; cDNA; 2511 BP.
 XX
 AC AAH33117;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen encoding cDNA SEQ ID NO:173.
 XX
 KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 7; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200122920-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US026524.
 XX
 PR 29-SEP-1999; 99US-0157137P.
 PR 03-NOV-1999; 99US-0163280P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX
 DR WPI; 2001-235357/24.
 DR P-PSDB; AAG73686.
 XX
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers.
 XX
 PS Claim 1; Page 2329-2330; 9803pp; English.
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
 CC proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene therapy
 CC and vaccine production. N and P may be used in the prevention, diagnosis
 CC and treatment of diseases associated with inappropriate P expression. For
 CC example, N and P may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of P by expressing inactive proteins or to
 CC supplement the patient's own production of P. Additionally, N may be used
 CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the proteins. N and P
 CC can be used in the prevention, diagnosis and treatment of colorectal
 CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent
 CC sequences used in the exemplification of the present invention. N.B.
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
 CC time of publication, meaning no sequences are present for SEQ ID NO:1027
 CC to 1052, 7921 and 7922
 XX
 SQ Sequence 2511 BP; 814 A; 445 C; 443 G; 803 T; 0 U; 6 Other;
 Query Match 15.8%; Score 584.2; DB 4; Length 2511;

Best Local Similarity 98.7%; Pred. No. 3.3e-126;		Matches 589; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	
QY	2763	CCAGCAAGCGGAACCTCTGGGAAAGCAATCTGTGATAAGGTCACTTCCCCCACTAAGG	2822
Db	73	CGAGCAAGCGGAACCTCTGGGAAAGCAATCTGTGATAAGGTCACTTCCCCCACTAAGG	132
QY	2823	TTTGAGACAGTTCCAGAAAGAACCCAGCTCAAGCGGAGCGAGCTCAGTTGTAGAGG	2882
Db	133	TTTGAGACAGTTCCAGAAAGAACCCAGCTCAAGCGGAGCGAGCTCAGTTGTAGAGG	192
QY	2883	GCTAATTCGCTCTGTTTGTATTTATGTTGATTTACTAAATTTGGGTTCAATATCTTTTAT	2942
Db	193	GCTAATTCGCTCTGTTTGTATTTATGTTGATTTACTAAATTTGGGTTCAATATCTTTTAT	252
QY	2943	TTTTCAATATCCCGAGTAAACCCAGTATATATCACTATATTTAATCAAGCTCTAGA	3002
Db	253	TTTTCAATATCCCGAGTAAACCCAGTATATATCACTATATTTAATCAAGCTCTAGA	312
QY	3003	GATGTTTCATGTAAGTAAAGTACTGCCCTTTGCACAGGAGCCTGTTTCTAAAGAAACCCATGCT	3062
Db	313	GATGTTTCATGTAAGTAAAGTACTGCCCTTTGCACAGGAGCCTGTTTCTAAAGAAACCCATGCT	372
QY	3063	GTGAAATAGAGACTTTTCTACTGATCATCAACTCTGTATCTGAGCAGTATACCAACC	3122
Db	373	GTGAAATAGAGACTTTTCTACTGATCATCAACTCTGTATCTGAGCAGTATACCAACC	432
QY	3123	ACATCTGAAGTCAACAGAGATCCCAAGTTTAAATTTGCCCTGGGAATGTGCGAGTATCT	3182
Db	433	ACATCTGAAGTCAACAGAGATCCCAAGTTTAAATTTGCCCTGGGAATGTGCGAGTATCT	492
QY	3183	AGAAAAATGAACCGTAGTTTTTTTGTATTTTAAATACAGAAAGTCATGTTTCTGCACTT	3242
Db	493	AGAAAAATGAACCGTAGTTTTTTTGTATTTTAAATACAGAAAGTCATGTTTCTGCACTT	552
QY	3243	TATAATTAAGCATGGAAGAAATTTCTTAGTAGGCAATTTGTACACTTTTGTGAAGTAAC	3302
Db	553	TATAATTAAGCATGGAAGAAATTTCTTAGTAGGCAATTTGTACACTTTTGTGAAGTAAC	612
QY	3303	CCATTTTCAGATTTGAAATACATCGGATAATGGTTGCTTTAAAAAAGAAAAAGATG	3359
Db	613	CCATTTTCAGATTTGAAATACATCGAATAATGGTTGCTTTAAAAAAGAAAAAGATG	669

Search completed: March 10, 2004, 11:00:07
Job time : 916.994 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:19:42 ; Search time 169.055 Seconds
(without alignments)
12152.440 Million cell updates/sec

Title: US-10-084-817-30

Perfect score: 3702

Sequence: 1 gcctagggaatgacaggcat.....aagggcggttaccgaggaac 3702

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
- 2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
- 3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
- 4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
- 5: /cgn2_6/prodata/2/ina/PCUS COMB.seq:*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1570.2	42.4	2215	1	US-08-285-440-14
2	1570.2	42.4	2215	1	US-08-630-349-14
3	1481.6	40.0	2137	1	US-08-285-440-13
4	1481.6	40.0	2137	1	US-08-630-349-13
5	1087.8	29.4	1677	1	US-08-285-440-12
6	1087.8	29.4	1677	1	US-08-630-349-12
7	999.2	27.0	1599	1	US-08-285-440-11
8	999.2	27.0	1599	1	US-08-630-349-11
9	937.4	25.3	939	1	US-08-285-440-10
10	937.4	25.3	939	1	US-08-630-349-10
11	428.6	11.6	606	3	US-09-385-982-295
12	350.8	9.5	369	1	US-08-285-440-9
13	350.8	9.5	369	1	US-08-630-349-9
14	350.4	9.5	357	1	US-08-285-440-8
15	350.4	9.5	357	1	US-08-630-349-8
16	346.4	9.4	348	1	US-08-285-440-7
17	346.4	9.4	348	1	US-08-630-349-7
18	264.2	7.1	273	1	US-08-285-440-26
19	264.2	7.1	273	1	US-08-630-349-26
20	263	7.1	285	1	US-08-285-440-23
21	263	7.1	285	1	US-08-630-349-23
22	173	4.7	7218	1	US-08-232-463-14
23	137.4	3.7	53526	3	US-08-658-136-2
24	137.4	3.7	53577	3	US-08-658-136-1
25	116.8	3.2	3489	2	US-08-728-323A-1
26	116.8	3.2	3489	4	US-09-298-568-1
27	116.8	3.2	3489	4	US-09-410-399-1

c 28	116.8	3.2	32207	2	US-08-770-379-20	Sequence 20, Appl
c 29	116.8	3.2	32207	3	US-08-757-669A-20	Sequence 20, Appl
c 30	116.8	3.2	32207	4	US-09-230-371A-20	Sequence 20, Appl
c 31	114.2	3.1	5433	3	US-08-929-329-1	Sequence 1, Appl
c 32	112.6	3.0	6755	3	US-08-931-999-4	Sequence 4, Appl
c 33	102.8	2.8	1926	4	US-09-249-585A-2	Sequence 2, Appl
c 34	102.8	2.8	1926	4	US-09-410-399-3	Sequence 3, Appl
c 35	102.8	2.8	2580	3	US-09-050-863-2	Sequence 2, Appl
c 36	102.8	2.8	2580	4	US-09-359-081-2	Sequence 2, Appl
c 37	102.8	2.8	5452	2	US-09-130-114-1	Sequence 1, Appl
c 38	102.8	2.8	8705	4	US-09-647-344A-14	Sequence 14, Appl
c 39	102.8	2.8	9600	3	US-08-910-647-1	Sequence 1, Appl
c 40	102.8	2.8	9600	4	US-09-620-925-1	Sequence 15, Appl
c 41	102.8	2.8	10596	1	US-07-884-811-15	Sequence 15, Appl
c 42	102.8	2.8	10596	1	US-07-885-971-15	Sequence 15, Appl
c 43	102.8	2.8	10596	1	US-08-087-783A-15	Sequence 15, Appl
c 44	102.8	2.8	10596	1	US-08-194-088B-15	Sequence 15, Appl
c 45	102.8	2.8	10596	2	US-08-194-087-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-285-440-14
; Sequence 14, Application US/08285440
; Patent No. 5532337
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/285,440
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/858,947
; FILING DATE: March 27, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2215 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE: Human
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:

3101 TATCTGACAGTGTATACCAACACATCTGAAGTCAACAGAGATCCAAAGTTTAAATTCG 3160
1959 TATCTGACAGTGTATACCAACACATCTGAAGTCAACAGAGATCCAAAGTTTAAATTCG 2018
3161 CTGGGAATGTGTGAGTATCTAGAAAATGAACCGTAGTTTGTGTTTAAATACAG 3220
2019 CTGGGAATGTGTGAGTATCTAGAAAATGAACCGTAG-TTTTGTGTTTAAATACAG 2077
3221 AAGTCATGTTGTTCTGCACTTATATATAAGCATGGAAGAAATTAATCTTAGTAGGCAAT 3280
2078 AAGTCATGTTGTTCTGCACTTATATATAAGCATGGAAGAAATTAATCTTAGTAGGCAAT 2137
3281 TGTAACATCTTTTGAAGTAACCATTTCCAGATTTCGAATATGCGATATGTTGCTTT 3340
2138 TGTAACATCTTTTGAAGTAACCATTTCCAGATTTCGAATATGCGATATGTTGCTTT 2197
3341 AAAAAAAGAAAAAGA 3357
2198 TAAAAAAGAAAAAGA 2214

RESULT 2
US-08-630-349-14
; Sequence 14, Application US/08630349
; Patent No. 5739008
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,349
; FILING DATE: April 10, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/285,440
; FILING DATE: August 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/858,947
; FILING DATE: March 27, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2215 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE: Human
; IMMEDIATE SOURCE: HeLa Cell
; FEATURE:
; NAME/KEY:
; LOCATION: 12-1695
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "human caldesmon of higher
; OTHER INFORMATION: molecular weight"
; FEATURE:
; NAME/KEY:
; LOCATION: 2103-2108
; IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "poly A signal"
; FEATURE:
; NAME/KEY:
; LOCATION: 2099-2215
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "poly A site"
US-08-630-349-14
Query Match 42.4%; Score 1570.2; DB 1; Length 2215;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 93; Indels 4; Gaps 3;
QY 1601 CAGAAACGTAAACAGCAGCTAGAAAGAGAAAAACATGCCATGCAAGAGACAAAGATAAAA 1660
DB 462 CAGAAAGATACAGATAGAGGAAACAGAAACAGTCCACCAAGTCTTACCAGAGAAATGAT 521
QY 1661 GGGGAAAAGGTAGAACAGAAAATAGAGGGAATGGTAAATGAAAAGAAAGACCAAGAA 1720
DB 522 TGGAGGGATGCTGAAGAAAAACAGAAAGAAAGCAAGGAAAAGGAGGAGGAGAGAGAG 581
QY 1721 GATAAACTTCAGACAGCTGCTCTTAAAGAAACAGGAGAGAGAGAGGAACTTAAAGTGCAA 1780
DB 582 AAGCCAAAGCGGGAGCATTGCGAGAAATTCAGGAGAGAGAGAGGAACTTAAAGTGCAA 641
QY 1781 GCTAAAAGAGAAAAGCTCCAAAGAACAGCCCTACCTTCAAAAAGAAAGAGATCAAAGAT 1840
DB 642 GCTAAAAGAGAAAAGCTCCAAAGAACAGCCCTACCTTCAAAAAGAAAGAGATCAAAGAT 701
QY 1841 GAAAAGATTAAAAAGGACAAAGAACCCAAAGAAAGAGTTAAGAGCTTTCATGGATCGAAAG 1900
DB 702 GAAAAGATTAAAAAGGACAAAGAACCCAAAGAAAGAGTTAAGAGCTTTCATGGATCGAAAG 761
QY 1901 AAGGGATTTACAGAACTTAAGTCGCGAATGGAAGAAATTCATGACCCCAAACTTAAACAT 1960
DB 762 AAGGGATTTACAGAACTTAAGTCGCGAATGGAAGAAATTCATGACCCCAAACTTAAACAT 821
QY 1961 ACTGAGAACTTTTTCAGCGCCCTCGAGGAGAGGCGCCAGCGTGGACACCAAGAGAGCTGAG 2020
DB 822 ACTGAGAACTTTTTCAGCGCCCTCGAGGAGAGGCGCCAGCGTGGACACCAAGAGAGCTGAG 881
QY 2021 GGGCGCCCCAGGTGAAAGCGCGCAAAAGGCTTGAGGAGCTTCGTCTGTCGCGCGGGAG 2080
DB 882 GGGCGCCCCAGGTGAAAGCGCGCAAAAGGCTTGAGGAGCTTCGTCTGTCGCGCGGGAG 941
QY 2081 ACCGAGAGCGAAGAGTTCCGAGAACTCAACAGAGCAGCAGGAGGCGGCTTTGGAGCTG 2140
DB 942 ACCGAGAGCGAAGAGTTCCGAGAACTCAACAGAGCAGCAGGAGGCGGCTTTGGAGCTG 1001
QY 2141 GAGGAACTCAAGAAAAAGAGGAGGAGAGAGAAAGGCTTCCTGGAGGAGGAGAGAGAGG 2200
DB 1002 GAGGAACTCAAGAAAAAGAGGAGGAGAGAGAAAGGCTTCCTGGAGGAGGAGAGAGAGG 1061
QY 2201 AGGAAGCAGGAGAAACCGATCCGAAACTCAGAGAGGAGGAGAGAGAGAGAGGCTTAAAG 2260
DB 1062 AGGAAGCAGGAGAAACCGATCCGAAACTCAGAGAGGAGGAGAGAGAGAGAGGCTTAAAG 1121
QY 2261 GAAGAGATTGAAAGGGGAT 2320
DB 1122 GAAGAGATTGAAAGGGGAT 1181
QY 2321 GCGTTGTCTAGATGACAGAAACCAATTCAGAGTTTCACTCTCTTAAGGTTTCACTCTCAAG 2380
DB 1182 GCGTTGTCTAGATGACAGAAACCAATTCAGAGTTTCACTCTCTTAAGGTTTCACTCTCAAG 1241
QY 2381 ATAGAAGAGCGAGCAGAAATTTTGAATTAAGTCTGTGAGAAAAAGCAGTGGTGTCAATCG 2440
DB 1242 ATAGAAGAGCGAGCAGAAATTTTGAATTAAGTCTGTGAGAAAAAGCAGTGGTGTCAATCG 1301
QY 2441 ACCCATCAAGCAGCAATAGTCTCCAGATTGACAGCAGACTGGAGCAGTATATACCAAGTCA 2500
DB 1302 ACCCATCAAGCAGCAATAGTCTCCAGATTGACAGCAGACTGGAGCAGTATATACCAAGTCA 1361
QY 2501 ATTGAGGGAACAAAAAGCGCAAAACCTTACAAAGCGGCGAGCCTTCGGATCTTCTCTTCT 2560

1362 ATTGAGGGAACAAAGGCGCAAAACCTACAAAGCGGCGAGCTCGATCTTCCGTCTCT 1421
 2561 GCTGAAGGTGTAGCCACATCAAGAGATATGCGGAGAAAGGAATGTGTTTTCATCCCC 2620
 1422 GCTGAAGGTGTAGCCACATCAAGAGATATGCGGAGAAAGGAATGTGTTTTCATCCCC 1481
 2621 ACTGCACAGGACACCAATAGGAAGTCTGCTGCTTGAAGGTAGGGTCTTAGCGGC 2680
 1482 ACTGCACAGGACACCAATAGGAAGTCTGCTGCTTGAAGGTAGGGTCTTAGCGGC 1541
 2681 ATCAATGAATGGTAACTAAACCCAGATGGAACCAAGTCCACCTGCTCCCAACCTTCT 2740
 1542 ATCAATGAATGGTAACTAAACCCAGATGGAACCAAGTCCACCTGCTCCCAACCTTCT 1601
 2741 GACTTGACAGCAGGACGTATCCAGCAGCGGACCTCTGGGAAAGCAATCTGTGAT 2800
 1602 GACTTGACAGCAGGACGTATCCAGCAGCGGACCTCTGGGAAAGCAATCTGTGAT 1661
 2801 AAGTCACTTCCCCCACTAAGGTTTGACACAGTTCAGAAAGAACCCCAAGCTCAAGAGCC 2860
 1662 AAGTCACTTCCCCCACTAAGGTTTGACACAGTTCAGAAAGAACCCCAAGCTCAAGAGCC 1721
 2861 AGACGAGCTCAGTTGAGAGGCTAATCGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2920
 1722 AGACGAGCTCAGTTGAGAGGCTAATCGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1779
 2921 AATTGGGTTCAATATCTTTTATTTTCAATATCCAGTAAACCCCATGATATATACATA 2980
 1780 AATTGGGTTCAATATCTTTTATTTTCAATATCCAGTAAACCCCATGATATATACATA 1839
 2981 TATTATAATACAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTGCGACAGAGCC 3040
 1840 TATTATAATACAG-CTAGAGATGTTTCATGTTAAAGTACTGCTTTGCGACAGAGCC 1898
 3041 TGTTCTAAAGAACCCCATGTTGAAATAGAGCTTTTCTACTGATCATATCACTCTG 3100
 1899 TGTTCTAAAGAACCCCATGTTGAAATAGAGCTTTTCTACTGATCATATCACTCTG 1958
 3101 TATCTGAGCAGTATACCAACACATCTGAAGTCAACAGATCCCAAGTTTAAATGTC 3160
 1959 TATCTGAGCAGTATACCAACACATCTGAAGTCAACAGATCCCAAGTTTAAATGTC 2018
 3161 CTGCGGATGTCAGTATCTAGAAATAGAACGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3220
 2019 CTGCGGATGTCAGTATCTAGAAATAGAACGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2077
 3221 AAGTCATGTTCTTCTGCACTTTATATAAAGCATGGAAGAAATATCTTAGTAGGCAAT 3280
 2078 AAGTCATGTTCTTCTGCACTTTATATAAAGCATGGAAGAAATATCTTAGTAGGCAAT 2137
 3281 TGTAACACTTTTGAAGTAAACCCATCTGATTTGATATCTGATATGTTGTTGTTGTTGTT 3340
 2138 TGTAACACTTTTGAAGTAAACCCATCTGATTTGATATCTGATATGTTGTTGTTGTTGTT 2197
 3341 AAAAAAAAAAGAAAAA 3357
 2198 TAAAAAAAAAAAAAAAAA 2214

RESULT 3

US-08-285-440-13
 ; Sequence 13, Application US/08285440
 ; Patent No. 5532337
 ; GENERAL INFORMATION:
 ; APPLICANT: Ken'ichiro HAYASHI et al.
 ; TITLE OF INVENTION: POLYPEPTIDE
 ; NUMBER OF SEQUENCES: 32
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wenderoth, Lind & Ponack
 ; STREET: 805 Fifteenth Street, N.W., #700
 ; CITY: Washington
 ; STATE: D.C.

COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: DisplayWrite
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/285,440
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/858,947
 FILING DATE: March 27, 1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Warren M. Cheek, Jr.
 REGISTRATION NUMBER: 33,367
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-8850
 TELEFAX:
 TELEX:
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2137 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cdna to mRNA
 HYPOTHETICAL:
 ANTI-SENSE:
 FRAGMENT TYPE:
 ORIGINAL SOURCE: Human
 ORGANISM:
 STRAIN:
 INDIVIDUAL ISOLATE:
 DEVELOPMENTAL STAGE:
 HAPLOTYPE:
 TISSUE TYPE:
 CELL TYPE:
 CELL LINE:
 ORGANELLE:
 IMMEDIATE SOURCE: Hela Cell
 LIBRARY:
 CLONE:
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT:
 MAP POSITION:
 UNITS:
 FEATURE:
 NAME/KEY:
 LOCATION: 12-1607
 IDENTIFICATION METHOD:
 OTHER INFORMATION: /note= "human caldesmon of lower
 OTHER INFORMATION: molecular weight"
 FEATURE:
 NAME/KEY:
 LOCATION: 2025-2030
 IDENTIFICATION METHOD:
 OTHER INFORMATION: /note= "poly A signal"
 FEATURE:
 NAME/KEY:
 LOCATION: 2121-2137
 IDENTIFICATION METHOD:
 OTHER INFORMATION: /note= "poly A site"
 PUBLICATION INFORMATION:
 AUTHORS:
 TITLE:
 JOURNAL:
 VOLUME:
 ISSUE:
 PAGES:
 DATE:

[illegible]

2712	GAACAAGTCACTGCTCCCAACCTTCTGACTTGGAGACGAGAGCGTATCCAGCAAGC	2771
1495	GAACAAGTCACTGCTCCCAACCTTCTGACTTGGAGACGAGAGCGTATCCAGCAAGC	1554
2772	GGAACTCTCGGAAAGCAATCTGTGGATAAGGTCACCTCCCCCACTAAGGTTTGAGACA	2831
1555	GGAACTCTCGGAAAGCAATCTGTGGATAAGGTCACCTCCCCCACTAAGGTTTGAGACA	1614
2832	GTTCCAGAAAGCCCAAGCTCAAGACGAGGACGAGCTCAGTGTGAGAGGCTAATTCG	2891
1615	GTTCCAGAAAGCCCAAGCTCAAGACGAGGACGAGCTCAGTGTGAGAGGCTAATTCG	1674
2892	CTCTGTTTGTATTTATGTTGATTTTACTAAATTTGGGTTTCATTTCTTTTCAATA	2951
1675	CTC--TTTGTATTTATGTTGATTTTACTAAATTTGGGTTTCATTTCTTTTCAATA	1732
2952	TCCCAGTAAACCCATGATATTTATCTATATTAATCAACAG-CTAGAGATGTTCAAT	3011
1733	TCCCAGTAAACCCATGATATTTATCTATATTAATCAACAG-CTAGAGATGTTCAAT	1791
3012	GGTAAAGTACTGCTCTTGCACAGGAGCCCTGTTTCTTAAAGAAACCCATGCTGGAATAG	3071
1792	GGTAAAGTACTGCTCTTGCACAGGAGCCCTGTTTCTTAAAGAAACCCATGCTGGAATAG	1851
3072	AGACTTTTCTACTGATCATATACTCTGTATCTGAGCAGTGATACCAACCATCTGAA	3131
1852	AGACTTTTCTACTGATCATATACTCTGTATCTGAGCAGTGATACCAACCATCTGAA	1911
3132	GTCAACAGAGATCCAAGTTTAAATTTGCTCGGGAATGTGTCAGTATCTAGAAAAATG	3191
1912	GTCAACAGAGATCCAAGTTTAAATTTGCTCGGGAATGTGTCAGTATCTAGAAAAATG	1971
3192	AACCGTAG-TTTTGTATTTTAAATACAGAGTCATGTTTCTGCACTTTATAATAA	3251
1972	AACCGTAG-TTTTGTATTTTAAATACAGAGTCATGTTTCTGCACTTTATAATAA	2030
3252	GCATGGAGAAATTTCTTAGTAGGCAATTTGTAACACTTTTTGAAAGTAACCCATTTTCAG	3311
2031	GCATGGAGAAATTTCTTAGTAGGCAATTTGTAACACTTTTTGAAAGTAACCCATTTTCAG	2090
3312	ATTGGAATCTCGCAATATGTTTGTCTTTAAAAAAGAAAAAGA	3357
2091	ATTGGAATCTCGCAATATGTTTGTCTTTAAAAAAGAAAAAGA	2136

RESULT 4
US-08-630-349-13
; Sequence 13, Application US/08630349
; Patent No. 5739008
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,349
; FILING DATE: April 10, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/285,440
; FILING DATE: August 4, 1994

```

1  RESULT 4
2  US-08-630-349-13
3  Sequence 13, Application US/08630349
4  Patent No. 5739008
5  GENERAL INFORMATION:
6  APPLICANT: Ken'ichiro HAYASHI et al
7  TITLE OF INVENTION: POLYPEPTIDE
8  NUMBER OF SEQUENCES: 32
9  CORRESPONDENCE ADDRESS:
10 ADDRESSEE: Wenderoth, Lind & Ponack
11 STREET: 905 Fifteenth Street, N.W.
12 CITY: Washington
13 STATE: D.C.
14 COUNTRY: U.S.A.
15 ZIP: 20005
16 COMPUTER READABLE FORM:
17 MEDIUM TYPE: Diskette, 3.5 inch, 1
18 COMPUTER: IBM Compatible
19 OPERATING SYSTEM: MS-DOS
20 SOFTWARE: Wordperfect 5.1
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: US/08/630,349
23 FILING DATE: April 10, 1996
24 CLASSIFICATION: 435
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: 08/285,440
27 FILING DATE: August 4, 1994

```


PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/858,947
 FILING DATE: March 27, 1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Warren M. Cheek, Jr.
 REGISTRATION NUMBER: 33,367
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2137 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 ORIGINAL SOURCE: Human
 IMMEDIATE SOURCE: HeLa Cell
 FEATURE:
 NAME/KEY: LOCATION: 12-1607
 IDENTIFICATION METHOD: /note= "human caldesmon of lower
 OTHER INFORMATION: molecular weight"
 FEATURE:
 NAME/KEY: LOCATION: 2025-2030
 IDENTIFICATION METHOD: /note= "poly A signal"
 OTHER INFORMATION:
 FEATURE:
 NAME/KEY: LOCATION: 2121-2137
 IDENTIFICATION METHOD:
 OTHER INFORMATION: /note= "poly A site"
 US-08-630-349-13

Query Match 40.0%; Score 1481.6; DB 1; Length 2137;
 Best Local Similarity 96.4%; Pred. No. 0;
 Matches 1548; Conservative 0; Mismatches 54; Indels 4; Gaps 3;

QY	1752	AGGGAAGAAAGAGGAACTAAAGTCAAGTAAAGAGAAAGCTCCAGAAAGCAAGC	1811
DB	535	AAGAAACCAAGAAAGAAAGCAAGGAAAGGAGGAGGAGAGAGCAAGCAAGCGAG	594
QY	1812	CTACCTTCAAAAGAGAGATCAAGATGAAGATTAAGAAAGGACAAAGCAAGCAAG	1871
DB	595	GGAGCATGGAGAAATCAGATCAAGATGAAGATTAAGAAAGGACAAAGCAAGCAAG	654
QY	1872	AAGAAGTTAAGAGCTTCATGATCGAAAGAGGAGTTTACAGAAAGTTAAGTCGCAAGATG	1931
DB	655	AAGAAGTTAAGAGCTTCATGATCGAAAGAGGAGTTTACAGAAAGTTAAGTCGCAAGATG	714
QY	1932	GAGAATTCATGACCCCAACAACTTAAACATACATGAGATACCTTACAGCCGCTGGAGGA	1991
DB	715	GAGAATTCATGACCCCAACAACTTAAACATACATGAGATACCTTACAGCCGCTGGAGGA	774
QY	1992	GGCCGACGCTGGACCAAGAGAGCTGAGGCGGCCCGCCAGGTGGAAGCCGCAAAAGGC	2051
DB	775	GGCCGACGCTGGACCAAGAGAGCTGAGGCGGCCCGCCAGGTGGAAGCCGCAAAAGGC	834
QY	2052	TGAGAGAGCTTCGTGCTGTCGCGGGAGAGCCGAGAGCAAGAGTTTCAGAAAGCTCAAC	2111
DB	835	TGAGAGAGCTTCGTGCTGTCGCGGGAGAGCCGAGAGCAAGAGTTTCAGAAAGCTCAAC	894
QY	2112	AGAAGCAGCAGGAGGCGGCTTGGAGCTGGAGGAACTCAAGAAAGAGGAGGAGAGAA	2171
DB	895	AGAAGCAGCAGGAGGCGGCTTGGAGCTGGAGGAACTCAAGAAAGAGGAGGAGAGAA	954
QY	2172	GGAAGTCTGGAGAGGAGAGAGCAGAGAGGAGAGCAGAGAGCCGATCGAAAGCTCA	2231
DB	955	GGAAGTCTGGAGAGGAGAGAGCAGAGAGGAGAGCAGAGAGCCGATCGAAAGCTCA	1014
QY	2232	GAGAGGAGGAGAGAGAGAGGAGGCTTAAAGGAGAGATTAAGGAGGAGAGAGAGCTG	2291
DB	1015	GAGAGGAGGAGAGAGAGAGGAGGCTTAAAGGAGAGATTAAGGAGGAGAGAGAGCTG	1074

QY	2292	CTGAGAAACGCCAGAAAGATGCCAGAAAGATGGCTTGTGAGATGACAGAAACCACTCAAGT	2351
DB	1075	CTGAGAAACGCCAGAAAGATGCCAGAAAGATGGCTTGTGAGATGACAGAAACCACTCAAGT	1134
QY	2352	GTTTCACTCTCTAAAGGTTTCATCTCTCAAGATAGAGAGCGAGCAGAAATTTTGAATAAGT	2411
DB	1135	GTTTCACTCTCTAAAGGTTTCATCTCTCAAGATAGAGAGCGAGCAGAAATTTTGAATAAGT	1194
QY	2412	CTGTGAGAAAGCAGTGGTGTCAATTCGACCCATCAAGCAGCAGTATGTCTCCAAGATTG	2471
DB	1195	CTGTGAGAAAGCAGTGGTGTCAATTCGACCCATCAAGCAGCAGTATGTCTCCAAGATTG	1254
QY	2472	ACAGCAGACTGGAGCAGTATACCAAGTCAATTCGAGGAAACAAAGCCGCAAAACCTACAA	2531
DB	1255	ACAGCAGACTGGAGCAGTATACCAAGTCAATTCGAGGAAACAAAGCCGCAAAACCTACAA	1314
QY	2532	AGCCGCGAGCTCGGATCTTCTGTTCTGCTGAAGGTGTACGCAATCAAGAGTATGT	2591
DB	1315	AGCCGCGAGCTCGGATCTTCTGTTCTGCTGAAGGTGTACGCAATCAAGAGTATGT	1374
QY	2592	GGGAGAAAGGAGTGTGTTTTCATCCCTGAGCAGGAGCAGCAGCAATTAAGGAAACTG	2651
DB	1375	GGGAGAAAGGAGTGTGTTTTCATCCCTGAGCAGGAGCAGCAGCAATTAAGGAAACTG	1434
QY	2652	CTGGCTTGAAGGTAGGGGTTTCTAGCCGATCAATGAATGCTTAACATAAACCAGATG	2711
DB	1435	CTGGCTTGAAGGTAGGGGTTTCTAGCCGATCAATGAATGCTTAACATAAACCAGATG	1494
QY	2712	GAAACAGTCACTGCTCCCAACCTTCTGACTTGAAGCAGGAGAGCAGTATCCAGCAAGC	2771
DB	1495	GAAACAGTCACTGCTCCCAACCTTCTGACTTGAAGCAGGAGAGCAGTATCCAGCAAGC	1554
QY	2772	GGAACTCTGGGAAAGCAATCTGTGATGAAGTCACTTCCCCCACTAAGGTTTGAGACA	2831
DB	1555	GGAACTCTGGGAAAGCAATCTGTGATGAAGTCACTTCCCCCACTAAGGTTTGAGACA	1614
QY	2832	GTTCCAGAAAGACCCAGCTCAAGACGAGGAGAGCTCAGTCTAGAGGCTTAATTCG	2891
DB	1615	GTTCCAGAAAGACCCAGCTCAAGACGAGGAGAGCTCAGTCTAGAGGCTTAATTCG	1674
QY	2892	CTCTGTTTGTATTTATGTTGATTTACTAAATGGGTTTCAATATCTTTTATTTTCAATA	2951
DB	1675	CTC--TTTGTATTTATGTTGATTTACTAAATGGGTTTCAATATCTTTTATTTTCAATA	1732
QY	2952	TCCAGTAAACCCATGTATATATCACTATATTTAATTAATCAGCTCTAGAGATGTTAT	3011
DB	1733	TCCAGTAAACCCATGTATATATCACTATATTTAATTAATCAGCTCTAGAGATGTTAT	1791
QY	3012	GGTAAAGTACTGCTTGTGACAGGAGCTGTTTCTAAAGAAACCCATGCTGTGAATAG	3071
DB	1792	GGTAAAGTACTGCTTGTGACAGGAGCTGTTTCTAAAGAAACCCATGCTGTGAATAG	1851
QY	3072	AGACTTTTCTGATCATCATTAATCTGATCTGAGCAGTGTATACCAACCACTGTAA	3131
DB	1852	AGACTTTTCTGATCATCATTAATCTGATCTGAGCAGTGTATACCAACCACTGTAA	1911
QY	3132	GTCAACAGAGATCCCAAGTTTAAATTTGCTGCGGAAATGTGTGAGTATCTAGAAAAATG	3191
DB	1912	GTCAACAGAGATCCCAAGTTTAAATTTGCTGCGGAAATGTGTGAGTATCTAGAAAAATG	1971
QY	3192	AACCGTAGTGTGTTTTTAAATACAGAAAGTCAATGTTGTTTCTGCACTTTATATAAA	3251
DB	1972	AACCGTAGTGTGTTTTTAAATACAGAAAGTCAATGTTGTTTCTGCACTTTATATAAA	2030
QY	3252	GCATGGAGAAATATCTTAGTAGGCAATTCGACACTTTTGAAGTAACCCATTTTCAG	3311
DB	2031	GCATGGAGAAATATCTTAGTAGGCAATTCGACACTTTTGAAGTAACCCATTTTCAG	2090
QY	3312	ATTGGAATACTGCGATAATGGTTGCTTTTAAAAAAGAAAAAGAA 3357	
DB	2091	ATTGGAATACTGCGATAATGGTTGCTTTTAAAAAAGAAAAAGAA 2136	

Db 1351 ATTGAGGGAACAAAAGCGCAAAACCTTACAAAGCCGAGCCTCGGATCTTCTCTGTTCT 1410
Qy 2561 GCTGAAGGTGTACGCAACATCAAGAGTATGTGGAGAAAGGAAATGTGTTTTCATCCCC 2620
Db 1411 GCTGAAGGTGTACGCAACATCAAGAGTATGTGGAGAAAGGAAATGTGTTTTCATCCCC 1470
Qy 2621 ACTGAGCGAGGACACCAATAAGGAACCTGCTGGCTTGAAGGTAGGGTTTCTAGCCGC 2680
Db 1471 ACTGAGCGAGGACACCAATAAGGAACCTGCTGGCTTGAAGGTAGGGTTTCTAGCCGC 1530
Qy 2681 ATCAATGAATGGCTAACTAAAAACCCAGATGGAACCAAGTCACCTGCTCCCAAAACCTTCT 2740
Db 1531 ATCAATGAATGGCTAACTAAAAACCCAGATGGAACCAAGTCACCTGCTCCCAAAACCTTCT 1590
Qy 2741 GACTTGAGACGAGAGACGTATCCAGACGAGCGGAACCTCTGGAAAGCAATCTGTGGAT 2800
Db 1591 GACTTGAGACGAGAGACGTATCCAGACGAGCGGAACCTCTGGAAAGCAATCTGTGGAT 1650
Qy 2801 AAGGTCACTTCCCCCACTAAGGTTTGA 2827
Db 1651 AAGGTCACTTCCCCCACTAAGGTTTGA 1677

RESULT 6

US-08-630-349-12
; Sequence 12, Application US/08630349
; Patent No. 5735008
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,349
; FILING DATE: April 10, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/285,440
; FILING DATE: August 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/858,947
; FILING DATE: March 27, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1677 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE: Human
; IMMEDIATE SOURCE: Hela Cell
; US-08-630-349-12
Query Match 29.4%; Score 1087.6; DB 1; Length 1677;
Best Local Similarity 92.9%; Pred. No. 1.7e-255;
Matches 1140; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
Qy 1601 CAGAAACGTAAACAGCAGCTAGAGAAAAAACAATGCCATGCAAGAGACAAAGATAAAA 1660

Db 451 CAAAGAAAGATACAGATAGAGGAAACAGAAACAGTCCACCAAGTCTCTACCAAGAAATGAT 510
Qy 1661 GGGGAAAAAGGTAGAACAGAAAAATAGAACGGGAATGGGTAAATGAATAAGAAAGAACCAAGAA 1720
Db 511 TGGAGGGATGCTGAAAGAAAAACAAGAAAGAGACAAGGAAAGGAGGAGGAGGAGGAG 570
Qy 1721 GATTAACCTTCAGACAGCTGCTCTAAAGAAAAACAGGAGAGAGAGAGGAACTTAAAGTGCNA 1780
Db 571 AAGCCAAAGCGGAGGAGCATTCGAGAAAAATCAAGGAGAGAGAGGGAACATAAGTGCNA 630
Qy 1781 GCTAAAAGAGAAAGCTCCAAAGAAAGCAAGCTTACCTTCAAAAAAAGAGAGATCAAAAGT 1840
Db 631 GCTAAAAGAGAAAGCTCCAAAGAAAGCAAGCTTACCTTCAAAAAAAGAGAGATCAAAAGT 690
Qy 1841 GAAAGATTTAAAAG 1900
Db 691 GAAAGATTTAAAAG 750
Qy 1901 AAGGGATTTACAGAAAGTTAAGTCGAGAAATGGAGAAATTCATGACCCCAAACTTAAACAT 1960
Db 751 AAGGGATTTACAGAAAGTTAAGTCGAGAAATGGAGAAATTCATGACCCCAAACTTAAACAT 810
Qy 1961 ACTGAGAAATACCTTTCAGCCGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2020
Db 811 ACTGAGAAATACCTTTCAGCCGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 870
Qy 2021 GGGGCCCCCAGGTGGAAAGCGGCAAAAGGCTTGGAGGAGCTTCGTCGTCGTCGTCGTCGTCGTCG 2080
Db 871 GGGGCCCCCAGGTGGAAAGCGGCAAAAGGCTTGGAGGAGCTTCGTCGTCGTCGTCGTCGTCGTCG 930
Qy 2081 ACCGAGAGCGAAGAGTTTCGAGAGCTCAACAGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 2140
Db 931 ACCGAGAGCGAAGAGTTTCGAGAGCTCAACAGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 990
Qy 2141 GAGGAACCTCAAGAAAAAGAGGAGAGAGAGAGAGAGAGGTCCTGGAGGAGAGAGAGAGAGAGAG 2200
Db 991 GAGGAACCTCAAGAAAAAGAGGAGAGAGAGAGAGAGAGGTCCTGGAGGAGAGAGAGAGAGAGAG 1050
Qy 2201 AGGAAGAGAGAGAGAGCGATTCGAAACTCAGAGAGGAGGAGAGAGAGAGAGAGGAGGAGGAGGAG 2260
Db 1051 AGGAAGAGAGAGAGAGCGATTCGAAACTCAGAGAGGAGGAGAGAGAGAGAGAGGAGGAGGAGGAG 1110
Qy 2261 GAAGAGATTCGAAGCGAGAGAGAGAGAGAGCTGCTGAGAAACCGCAGAGAGATGCCAGAGAGAT 2320
Db 1111 GAAGAGATTCGAAGCGAGAGAGAGAGAGCTGCTGAGAAACCGCAGAGAGATGCCAGAGAGAT 1170
Qy 2321 GGCTTGTGATGACAAAGAAACCAATTCAGGTGTTTCACTCTTAAAGGTTTCACTCTCTCAAG 2380
Db 1171 GGCTTGTGATGACAAAGAAACCAATTCAGGTGTTTCACTCTTAAAGGTTTCACTCTCTCAAG 1230
Qy 2381 ATAGAGAGCGAGCAGAGAAATTTTGAATAAGTCTGTCAGAAAAAGCAGTGGTGTCAAAATCG 2440
Db 1231 ATAGAGAGCGAGCAGAGAAATTTTGAATAAGTCTGTCAGAAAAAGCAGTGGTGTCAAAATCG 1290
Qy 2441 ACCCATCAAGCAGCAATAGTCTCCAAGATTGACAGAGCTGGAGCAGTATACAGTGCA 2500
Db 1291 ACCCATCAAGCAGCAATAGTCTCCAAGATTGACAGAGCTGGAGCAGTATACAGTGCA 1350
Qy 2501 ATTGAGGAGCAAAAAGCGCAAAACCTCAAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2560
Db 1351 ATTGAGGAGCAAAAAGCGCAAAACCTCAAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1410
Qy 2561 GCTGAAGGTGTACGCAACATCAAGAGTATGTGGAGAAAGGGAATGTGTTTTCATCCCC 2620
Db 1411 GCTGAAGGTGTACGCAACATCAAGAGTATGTGGAGAAAGGGAATGTGTTTTCATCCCC 1470
Qy 2621 ACTGAGGAGCAGCAGCAACCAATAAGGAACCTGCTGGCTTGAAGGTAGGGTTTCTAGCCGC 2680
Db 1471 ACTGAGGAGCAGCAGCAACCAATAAGGAACCTGCTGGCTTGAAGGTAGGGTTTCTAGCCGC 1530
Qy 2681 ATCAATGAATGGCTAACTAAAAACCCAGATGGAACCAAGTCACCTGCTCCCAAAACCTTCT 2740

1531	ATCAATGAATGGCTAACTATAAAACCCAGATGGGAAAACAAGTCACCTGCTCCCAAACTTCT	1590
2741	GACTTTGAGACCGAGGAGACGTATCCAGCAAGCGGAACCTCTGGGAAAAGCAATCTGTGGAT	2800
1591	GACTTTGAGACCGAGGAGACGTATCCAGCAAGCGGAACCTCTGGGAAAAGCAATCTGTGGAT	1650
2801	AAGGTCACATTCCCCCACTAAGGTTTGA	2827
1651	AAGGTCACATTCCCCCACTAAGGTTTGA	1677

RESULT 7

JS-08-285-440-11
Sequence 11, Application US/08285440
Patent No. 5532337
GENERAL INFORMATION:
APPLICANT: Ken'ichiro HAYASHI et al.
TITLE OF INVENTION: POLYPEPTIDE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 KB
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/285,440
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/858,947
FILING DATE: March 27, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1599 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE: Human
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE: Hela Cell
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:

Sequence 10, Application US/08285440
 Patent No. 5532337
 GENERAL INFORMATION:
 APPLICANT: Ken'ichiro HAYASHI et al.
 TITLE OF INVENTION: POLYPEPTIDE
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wenderoth, Lind & Ponack
 STREET: 805 Fifteenth Street, N.W., #700
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: DisplayWrite
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/285,440
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/858,947
 FILING DATE: March 27, 1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Warren M. Cheek, Jr.
 REGISTRATION NUMBER: 33,367
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-8850
 TELEFAX:
 TELEX:
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 939 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL:
 ANTI-SENSE:
 FRAGMENT TYPE:
 ORIGINAL SOURCE: Human
 ORGANISM:
 STRAIN:
 INDIVIDUAL ISOLATE:
 DEVELOPMENTAL STAGE:
 HAPLOTYPE:
 TISSUE TYPE:
 CELL TYPE:
 CELL LINE:
 ORGANELLER:
 IMMEDIATE SOURCE: HeLa Cell
 LIBRARY:
 CLONE:
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT:
 MAP POSITION:
 UNITS:
 FEATURE:
 NAME/KEY:
 LOCATION:
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 PUBLICATION INFORMATION:
 AUTHORS:
 TITLE:
 JOURNAL:
 VOLUME:
 ISSUE:
 PAGES:
 DATE:

DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO:
 US-08-285-440-10
 Query Match 25.3%; Score 937.4; DB 1; Length 939;
 Best Local Similarity 99.9%; Pred. No. 6.1e-219;
 Matches 938; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1889 ATGGATCGAAAGAGGGATTACAGAAAGTTAACTCGCAGAAATTCAGAAATTCATGACCCAC 1948
 DB 1 ATGGATCGAAAGAGGGATTACAGAAAGTTAACTCGCAGAAATTCAGAAATTCATGACCCAC 60
 QY 1949 AAACCTAAACATACCTAGAAATCTTTACGCGCCCTCGAGAGGGCCAGCGTGGACACC 2008
 DB 61 AAACCTAAACATACCTAGAAATCTTTACGCGCCCTCGAGAGGGCCAGCGTGGACACC 120
 QY 2009 AAGGAGGCTGAGGGGCGCCCGCCAGGTGGAAGCGGCAAAAGGCTGGAGGAGCTTCGTCGT 2068
 DB 121 AAGGAGGCTGAGGGGCGCCCGCCAGGTGGAAGCGGCAAAAGGCTGGAGGAGCTTCGTCGT 180
 QY 2069 CGTCGCGGGAGACCGAGAGCGAAGTTTCGAGAAAGCTCAAAACAGAGCAGCAGAGGCG 2128
 DB 181 CGTCGCGGGAGACCGAGAGCGAAGTTTCGAGAAAGCTCAAAACAGAGCAGCAGAGGCG 240
 QY 2129 GCTTTGGAGCTGAGGAACTCAAGAAAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAG 2188
 DB 241 GCTTTGGAGCTGAGGAACTCAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
 QY 2189 GAAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2248
 DB 301 GAAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
 QY 2249 AGGAGGCTAAAGGAGAGAGATTGAAAGCGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2308
 DB 361 AGGAGGCTAAAGGAGAGAGATTGAAAGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
 QY 2309 ATGCCAGAGAGTGGCTTGTTCAGATGACAGAAACCATTTCAAGTGTTCCTCACTTCAAGGT 2368
 DB 421 ATGCCAGAGAGTGGCTTGTTCAGATGACAGAAACCATTTCAAGTGTTCCTCACTTCAAGGT 480
 QY 2369 TCATCTCTCAAGATAGAGAGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2428
 DB 481 TCATCTCTCAAGATAGAGAGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
 QY 2429 GGTGTCAAAATCGAACCATCAAGCAGCAATAGTCTCAAGATTGACAGCAGACTGGAGCAG 2488
 DB 541 GGTGTCAAAATCGAACCATCAAGCAGCAATAGTCTCAAGATTGACAGCAGACTGGAGCAG 600
 QY 2489 TATACAGTGCAATTGAGGGACAAAAGCGCAAAACCTCAAAAGCGGCGGAGCTCGGAT 2548
 DB 601 TATACAGTGCAATTGAGGGACAAAAGCGCAAAACCTCAAAAGCGGCGGAGCTCGGAT 660
 QY 2549 CTTCTGTTCTCTGCTGAAAGGTGACCAACATCAAGAGTATGTGGGAGAAAGGGAATGTG 2608
 DB 661 CTTCTGTTCTCTGCTGAAAGGTGACCAACATCAAGAGTATGTGGGAGAAAGGGAATGTG 720
 QY 2609 TTTTATCCCCCAGCTGAGCAGGACACCAAAATAAGGAAACTGTGGTTGAAGTAGGG 2668
 DB 721 TTTTATCCCCCAGCTGAGCAGGACACCAAAATAAGGAAACTGTGGTTGAAGTAGGG 780
 QY 2669 GTTCTAGCGGCATCAATGATGGCTAACTAAACCCAGATGGAACCAAGTCACTGCT 2728
 DB 781 GTTCTAGCGGCATCAATGATGGCTAACTAAACCCAGATGGAACCAAGTCACTGCT 840
 QY 2729 CCAAAACCTTCTGACTTGAGACCGAGAGAGGATCCAGAGCGGAACTCTTGGGAAAAG 2788
 DB 841 CCAAAACCTTCTGACTTGAGACCGAGAGAGGATCCAGAGCGGAACTCTTGGGAAAAG 900
 QY 2789 CAATCTGTGGATAAGGTCACTTCCCCCAGTAAAGTTGA 2827
 DB 901 CAATCTGTGGATAAGGTCACTTCCCCCAGTAAAGTTGA 939

	QY	2249	AGGAGCTTAAGAAGACAGATTTCGAAGGCCAAGCAGACGATGGCTCGAGAAAACCCCAAGAAG	2300
	DB	361	AGGAGCTTAAGAAGACAGATTTCGAAGGCCAAGCAGACGATGGCTCGAGAAAACCCCAAGAAG	420
	QY	2309	ATGCCAAGAAGATGGCTTGTGCAGATGACAAGAAACCATTCAAAGTGTTTCACTCTCTAAAGGT	2368
	DB	421	ATGCCAAGAAGATGGCTTGTGCAGATGACAAGAAACCATTCAAAGTGTTTCACTCTCTAAAGGT	480
	QY	2369	TCACTCTCAAGATAGAACAGCGAGCAGAGAAATTTTTGAAATAAGTCTGTGCAGAAAAACGAGT	2428
	DB	481	TCACTCTCAAGATAGAACAGCGAGCAGAGAAATTTTTGAAATAAGTCTGTGCAGAAAAACGAGT	540
	QY	2429	GGTGTCAAAATCGACCCCATCAAGCAGCAATAGTCTCCAAGATTGACAGCAGACTGGAGCAG	2488
	DB	541	GGTGTCAAAATCGACCCCATCAAGCAGCAATAGTCTCCAAGATTGACAGCAGACTGGAGCAG	600
	QY	2489	TATACCAAGTGCATTCAGCGGNACAAAAGCGCAAAACCTTACAAAGCGCGCAGGCTCGGAT	2548
	DB	601	TATACCAAGTGCATTCAGCGGNACAAAAGCGCAAAACCTTACAAAGCGCGCAGGCTCGGAT	660
	QY	2549	CTTCCCTGTTCTCTGCTGAAGGTGTACCAACATCAAGAGTAGTGTGGAGAAAGGGAATGTG	2608
	DB	661	CTTCCCTGTTCTCTGCTGAAGGTGTACCAACATCAAGAGTAGTGTGGAGAAAGGGAATGTG	720
	QY	2609	TTTTTCATCCCCCACTCCAGCAGGCACACCAAAATAAGGAACACTCTGCTTTGAAAGGTAGGG	2668
	DB	721	TTTTTCATCCCCCACTCCAGCAGGCACACCAAAATAAGGAACACTCTGCTTTGAAAGGTAGGG	780
	QY	2669	GTTTCTAGCGCGCATCAATGAATGGCTAACTPAAACCCCAGATGGAAACAAGTCACTGCT	2728
	DB	781	GTTTCTAGCGCGCATCAATGAATGGCTAACTPAAACCCCAGATGGAAACAAGTCACTGCT	840
	QY	2729	CCCAAACTTTCAGTCTGACACCAAGCAGAGAGCGTATCCAGCAAGCGGAACCTCTCGGAAAAAG	2788
	DB	841	CCCAAACTTTCAGTCTGACACCAAGCAGAGAGCGTATCCAGCAAGCGGAACCTCTCGGAAAAAG	900
	QY	2789	CAATCTGTGGATAAGGTCACTTCCCCCACTAAAGTTTGA	2827
	DB	901	CAATCTGTGGATAAGGTCACTTCCCCCACTAAAGTTTGA	939
 RESULT 11				
US-09-385-982-295/c				
; Sequence 295, Application US/09385982				
; Patent No. 6262334				
; GENERAL INFORMATION:				
; APPLICANT: ENDEGE, WILSON O., ET AL.				
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION				
; FILE REFERENCE: CCDNA-260XX				
; CURRENT APPLICATION NUMBER: US/09/385,982				
; CURRENT FILING DATE: 1999-08-30				
; EARLIER APPLICATION NUMBER: 09/328,111				
; EARLIER FILING DATE: 1999-06-08				
; EARLIER APPLICATION NUMBER: 60/117,393				
; EARLIER FILING DATE: 1999-01-27				
; EARLIER APPLICATION NUMBER: 60/098,639				
; EARLIER FILING DATE: 1998-08-31				
; NUMBER OF SEQ ID NOS: 544				
; SOFTWARE: FastSeq for Windows Version 3.0				
; SEQ ID NO 295				
; LENGTH: 606				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
; FEATURE:				
; NAME/KEY: misc feature				
; LOCATION: (1)..(606)				
; OTHER INFORMATION: n = A,T,C or G				
US-09-385-982-295				

Query Match 11.6%; Score 428.6; DB 3; Length 606;
Best Local Similarity 96.5%; Pred. No. 6.1e-95;

```

Matches 437; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
2y 2570 GTAGCAACATCAAGAGTATGTGGGAGAAAGGAATGCTTTTCATCCCCCACTGCAGCA 2629
Db 457 GTAGCAACATCAAGAGTATGTGGGAGAAAGGAATGCTTTTCATCCCCCACTGCAGCA 398
2y 2630 GGCACACCAATAAGGAACTCTGCTTGAAGGTAGGGGTTTCTAGCCGCATCAATGAA 2689
Db 397 GGCACACCAATAAGGAACTCTGCTTGAAGGTAGGGGTTTCTAGCCGCATCAATGAA 338
2y 2690 TGGCTAATCAAAACCCAGATGGAACAGTCACTCTCCCAACCTCTGACTTGAGA 2749
Db 337 TGGCTAATCAAAACCCAGATGGAACAGTCACTCTCCCAACCTCTGACTTGAGA 278
2y 2750 CCAGGAGAGCTATCCAGCAAGCGGACCTCTGGGAAAGCAATCTGTGGATAAGGTCACT 2809
Db 277 CCAGGAGAGCTATCCAGCAAGCGGACCTCTGGGAAAGCAATCTGTGGATAAGGTCACT 218
2y 2810 TCCCCCACTAAGTTTGACACAGTTCCAGAAAGACCCAGCTCAAGACGACGACGAGC 2869
Db 217 TCCCCCACTAAGTTTGACACAGTTCCAGAAAGACCCAGCTCAAGACGACGACGAGC 158
2y 2870 TCAGTTGTAGAGGGTAAATCGCTCTGTTTGTATTATCTTGTATTATCTTGAATTTACTAAATTTGGTT 2929
Db 157 TCAGTTGTAGAGGGTAAATCGCTCTGTTTGTATTATCTTGTATTATCTTGAATTTACTAAATTTGGTT 98
2y 2930 CATATCTTTTATTTTCAATATCCAGTAACCCATGATATATTAATTAATTAATA 2989
Db 97 CATATCTTTTATTTTCAATATCCAGTAACCCATGATATATTAATTAATTAATA 38
2y 2990 ATCAGAGCTAGAGATGTTTCATGGTAAAGTAC 3022
Db 37 ATCAGAGCTAGAGATGTTTCATGGTAAAGTAC 5

```

RESULT 12

```

US-08-285-440-9
; Sequence 9, Application US/08285440
; Patent No. 5832337
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/285,440
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/858,947
; FILING DATE: March 27, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEICAL:
; ANTI-SENSE:
; FRAGMENT TYPE: Human
; ORIGINAL SOURCE: Human
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE: Hela Cell
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-08-285-440-9

```

```

Query Match 9.5%; Score 350.8; DB 1; Length 369;
Best Local Similarity 99.4%; Pred.No. 4.4e-76;
Matches 352; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2474 AGCAGACTGGAGCAGTATACCACTGCAATTTGAGGGAACAAAGGCGCAAAACCTTCAAG 2533
Db 16 AGCAGACTGGAGCAGTATACCACTGCAATTTGAGGGAACAAAGGCGCAAAACCTTCAAG 75
QY 2534 CCGGCGAGCCTCGGATCTTCTGTTTCTGCTGAAGGTGTACGCAACATCAAGAGTATGTGG 2593
Db 76 CCGGCGAGCCTCGGATCTTCTGTTTCTGCTGAAGGTGTACGCAACATCAAGAGTATGTGG 135
QY 2594 GAGAAAGGGAATGTGTTTTTCATCCCACTGCGAGGACACCAAAATAGGAAACTGTCT 2653
Db 136 GAGAAAGGGAATGTGTTTTTCATCCCACTGCGAGGACACCAAAATAGGAAACTGTCT 195
QY 2654 GGCTTGAAGTAGGGTTTCTTACCGCATCAATGAATGGCTAACTAAACCCAGATGA 2713
Db 196 GGCTTGAAGTAGGGTTTCTTACCGCATCAATGAATGGCTAACTAAACCCAGATGA 255
QY 2714' AACAAAGTCACTGCTCCCAACCTTCTGACTTTGAGACGAGAGCTATCCAGCAAGCGG 2773
Db 256 AACAAAGTCACTGCTCCCAACCTTCTGACTTTGAGACGAGAGCTATCCAGCAAGCGG 315
QY 2774 AACCTCTGGGAAAGCAATCTGTGATAAGTCACTTCCCACTAAAGTTTGA 2827
Db 316 AACCTCTGGGAAAGCAATCTGTGATAAGTCACTTCCCACTAAAGTTTGA 369

```

RESULT 13

Query Match	9.5%;	Score 350.8;	DB 1;	Length 369;
Best Local Similarity	99.4%;	Prod. No. 4.4e-76;		
Matches 352;	Conservative	Mismatches 2;	Indels 0;	Gaps 0;
QY	2474	ACGACGACTCGAGCAGTATACGACGTCGAATTCGAGGGAAACAAAAGCGCAAAACCTCAAAAG	2533	
DB	16	AGAGAGACTGAGCAGTATACAGCTGCAATTCGAGGGAAACAAAAGCGCAAAACCTCAAAAG	75	
QY	2534	CGGGCAGCCTTCGGATCTTCCTGCTGTAAGGTGTACGCAACATCAAGAGTATGTGG	2593	
DB	76	CGGGCAGCCTTCGGATCTTCCTGCTGTAAGGTGTACGCAACATCAAGAGTATGTGG	135	
QY	2594	GAGAAAGGGAATGTGTTTTTCATCCCCCACTGCAAGCAGGCACACCAAAATAAGGAAATCTGCT	2653	
DB	136	GAGAAAGGGAATGTGTTTTTCATCCCCCACTGCAAGCAGGCACACCAAAATAAGGAAATCTGCT	195	
QY	2654	GGCTTGAAGTACGGGTTTCTAGCGGCATCAATGAATGGCTAACTAAACCCCAAGATGGA	2713	
DB	196	GGCTTGAAGTACGGGTTTCTAGCGGTATCAATGAATGGCTAACTAAACCCCAAGATGGA	255	
QY	2714	AACAAGTCACCTGCTCCCAACCTTCTCACTTTGACACCAAGGAGACGTATCCAGCAAGCGG	2773	
DB	256	ARCAAGTCACCTGCTCCCAACCTTCTCACTTTGACACCAAGGAGACGTATCCAGCAAGCGG	315	
QY	2774	AACCTCTGGGAAAAGCAATCTGTGGATAAGGTCACCTCCCCCACTAAGCTTTGA	2827	
DB	316	AACCTCTGGGAAAAGCAATCTGTGGATAAGGTCACCTCCCCCACTAAGGTTTGA	369	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2004, 03:37:21 ; Search time 862.637 Seconds
(without alignments)
15799.545 Million cell updates/sec

Title: US-10-084-817-30
Perfect score: 3702
Sequence: 1 gctagggaatgacaggcat.....aaggggcggttacggagaac 3702

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 184079884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications NA:*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
 - 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
 - 13: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
 - 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
 - 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
 - 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
 - 17: /cgn2_6/ptodata/1/pubpna/US50_NEW_PUB.seq:*
 - 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3702	100.0	3702	14	US-10-084-817-30.
2	3201.2	86.5	3500	15	US-10-062-674-1815
3	3116.4	84.2	3610	9	US-09-880-107-2406
4	3116.4	84.2	3610	10	US-09-873-267C-721
5	2353	63.6	4586	15	US-10-062-674-1817
6	1793.6	48.4	4547	9	US-09-974-298-36
7	1793.6	48.4	4547	14	US-10-084-817-31
8	1771	47.8	4261	15	US-10-062-674-1816
9	1578	42.6	2845	15	US-10-341-434-136
10	1578	42.6	2845	15	US-10-341-434-215
11	1455.8	39.3	3342	9	US-09-765-231A-32
12	584.2	15.8	2511	9	US-09-925-301-100
13	584.2	15.8	2511	14	US-10-106-698-183
14	546.8	14.8	2209	9	US-09-974-298-35
15	499	13.5	541	12	US-10-085-783A-16172

16	499	13.5	541	15	US-10-242-535A-16172
17	495.8	13.4	543	14	US-10-198-846-9612
18	447.4	12.1	449	14	US-10-060-036-1194
19	436.6	11.8	503	10	US-09-918-995-19285
20	428.6	11.6	436	10	US-09-918-995-3028
21	428.6	11.6	606	10	US-09-871-161-295
22	412.8	11.2	2850	14	US-10-198-846-13694
23	284	7.7	232	9	US-09-234-093B-5212
24	229.6	6.2	339	9	US-09-925-301-689
25	228.4	6.2	432	15	US-10-264-049-1570
26	211.4	5.7	400	9	US-09-960-352-3498
27	186.4	5.0	480	12	US-10-085-783A-50761
28	186.4	5.0	480	15	US-10-242-535A-50761
29	165.4	4.5	720	13	US-10-101-487-76
30	164.2	4.4	1840	14	US-10-017-161-2359
31	164.2	4.4	1840	15	US-10-292-798-2003
32	164	4.4	720	13	US-10-101-487-74
33	148.6	4.0	169	12	US-10-085-783A-18790
34	148.6	4.0	169	15	US-10-242-535A-18790
35	145	3.9	533	10	US-09-918-624B-55
36	137.4	3.7	53522	10	US-09-904-968A-1
37	135.2	3.7	554	13	US-10-101-487-106
38	133	3.6	554	13	US-10-101-487-69
39	132.8	3.6	530	13	US-10-101-487-73
40	132.6	3.6	522	13	US-10-101-487-71
41	129.6	3.5	68233	15	US-10-034-650-31
42	124.8	3.4	6668	14	US-10-311-455-1670
43	124.4	3.4	42533	14	US-10-004-113-43
44	123.6	3.3	2479	15	US-10-108-260A-1966
45	122	3.3	903	14	US-10-029-386-20462

ALIGNMENTS

RESULT 1
US-10-084-817-30
; Sequence 30, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Pion
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 30
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 1312325CB1
US-10-084-817-30

Query Match 100.0%; Score 3702; DB 14; Length 3702;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3702; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTAGGAATGACAGGCATCTCCACAGGAGGTGTCATCCACCTTGGTGGGTGCTG 60
Db 1 GCCTAGGAATGACAGGCATCTCCACAGGAGGTGTCATCCACCTTGGTGGGTGCTG 60

QY 61 CATTGGCTGCTATTAGAAAAACACAGGACAATGCATACCCGCTCCCGACTGTAA 120
Db 61 CATTGGCTGCTATTAGAAAAACACAGGACAATGCATACCCGCTCCCGACTGTAA 120

Qy	121	CATAGGGATATGTGTTCACTTAGCATGCACTTCTGGAGGGGCAAGGAAGGGCGTCT	180
Db	121	CATAGGGATATGTGTTCACTTAGCATGCACTTCTGGAGGGGCAAGGAAGGGCGTCT	180
Qy	181	GGAGTTTTATTGAATAGAGCAGTGTGTATTTCGGCTGCTGCTGCCCGCTGCTTGTCTCT	240
Db	181	GGAGTTTTATTGAATAGAGCAGTGTGTATTTCGGCTGCTGCTGCCCGCTGCTTGTCTCT	240
Qy	241	CTGGCTGTGCTCCTGCTTAAAGAAATCAGTCCTTCCCTTCCGACTTAGTCTCGGAGAGA	300
Db	241	CTGGCTGTGCTCCTGCTTAAAGAAATCAGTCCTTCCCTTCCGACTTAGTCTCGGAGAGA	300
Qy	301	AGTTTCAGACTACAAGGTATCATTTGGAACATTTCAAGATCATCAAAATTCACAGG	360
Db	301	AGTTTCAGACTACAAGGTATCATTTGGAACATTTCAAGATCATCAAAATTCACAGG	360
Qy	361	GATTTGGTCGACCAACAGAGAGGCTCAGACATCTGATTTGCTGACCTGTCCAGACATCATCTG	420
Db	361	GATTTGGTCGACCAACAGAGAGGCTCAGACATCTGATTTGCTGACCTGTCCAGACATCATCTG	420
Qy	421	GTCTCCCTGAACTGAAATCACACCATGGATGATTTTGAAGCGTCCGAGAACTTAGAAG	480
Db	421	GTCTCCCTGAACTGAAATCACACCATGGATGATTTTGAAGCGTCCGAGAACTTAGAAG	480
Qy	481	GCAAAAGGGAGGAGATCGCATTCGAAGCAGAAAGAAATCGCTTACAGAGGAATGACGA	540
Db	481	GCAAAAGGGAGGAGATCGCATTCGAAGCAGAAAGAAATCGCTTACAGAGGAATGACGA	540
Qy	541	TGATGAAGAGGAGGACCGCGGAAACCGCGCCGCGAGCCCGACAGGAACGGCTCGGCA	600
Db	541	TGATGAAGAGGAGGACCGCGGAAACCGCGCCGCGAGCCCGACAGGAACGGCTCGGCA	600
Qy	601	GAAGCAGGAGGAAGATCTTTGGCAGAGTGACCGACCCAGGTGGAGGTGAATGCCAGAA	660
Db	601	GAAGCAGGAGGAAGATCTTTGGCAGAGTGACCGACCCAGGTGGAGGTGAATGCCAGAA	660
Qy	661	CAGTGTGCTGACGAGGAGCCCAAGACAACCAACAACACTCAAGTGGAAAGGGGATGA	720
Db	661	CAGTGTGCTGACGAGGAGCCCAAGACAACCAACAACACTCAAGTGGAAAGGGGATGA	720
Qy	721	TCAGGCGCGCATTTCTGGAGCGCTGGCTCGGCGTGAGGAAAGACGCCAAAACCGCTTCA	780
Db	721	TCAGGCGCGCATTTCTGGAGCGCTGGCTCGGCGTGAGGAAAGACGCCAAAACCGCTTCA	780
Qy	781	GGAGGCTCTGGAGCGGAGAGGATTCGACCCCAACAATACAGATGCAAGTCTGCGT	840
Db	781	GGAGGCTCTGGAGCGGAGAGGATTCGACCCCAACAATACAGATGCAAGTCTGCGT	840
Qy	841	CCCAAGCAGAAAGATGCAGAAATGACACAGCAGAAATAGAACTACCGAGAGGAGAA	900
Db	841	CCCAAGCAGAAAGATGCAGAAATGACACAGCAGAAATAGAACTACCGAGAGGAGAA	900
Qy	901	AAGTGAAAGTCCGCAAGAAAGATACGAGATAGAGGAAACAGAAACAGTCACCAGTCTTA	960
Db	901	AAGTGAAAGTCCGCAAGAAAGATACGAGATAGAGGAAACAGAAACAGTCACCAGTCTTA	960
Qy	961	CCAGAAAGATGATTGGAGGGATGCTGAAAGAAAACAAGAAAGACAGGAAAGGAGGA	1020
Db	961	CCAGAAAGATGATTGGAGGGATGCTGAAAGAAAACAAGAAAGACAGGAAAGGAGGA	1020
Qy	1021	GGAGGAAGGAGGAGAGCCAAAGCAGGAGGACATTCGAGAAATCAGGTAGAGTGAATGCT	1080
Db	1021	GGAGGAAGGAGGAGAGCCAAAGCAGGAGGACATTCGAGAAATCAGGTAGAGTGAATGCT	1080
Qy	1081	GGAAAGAGAAAACCAACTGAAAGCCAGGAGGAAACAGTGGTAAATGCTATTAATAAATCGGCA	1140
Db	1081	GGAAAGAGAAAACCAACTGAAAGCCAGGAGGAAACAGTGGTAAATGCTATTAATAAATCGGCA	1140
Qy	1141	GATCAGTCTCAGAGAGCCTTAAACAAGAGGAGGAGGAGGAGGATTCAGATGAGATTC	1200
Db	1141	GATCAGTCTCAGAGAGCCTTAAACAAGAGGAGGAGGAGGAGGATTCAGATGAGATTC	1200
Qy	1201	CCATCATGAAAAGATGGAAGAGGAAGACAAGGAAAGAGCTGAGGACAGAGGAGGCAAGGTT	1260

Db	1201	CCATCATGAAAGAATGGAACAGAGGAGCAGAGAAAGAGCTTGAGCGCAGAGAGGGCAGGTT	1260
Qy	1261	GGAGCGAAGAAAGAGAGAAAGAAATTTAAAGCCGAGCAAGACAAAAGATAGCAGATGAACG	1320
Db	1261	GGAAAGCAAGAAAGAGAGAAAGAAATTTAAAGCCGAGCAAGACAAAAGATAGCAGATGAACG	1320
Qy	1321	AGCAAGAAATTTGAAGCAGAGAAAGAAAGCAGCTGCCCCAAGAAAGACAAAAGAGAGAGGCAGA	1380
Db	1321	AGCAAGAAATTTGAAGCAGAGAAAGAAAGCAGCTGCCCCAAGAAAGACAAAAGAGAGAGGCAGA	1380
Qy	1381	AGAGAGGAAAGGATGAGGGAGAAAGAGAAAGAGGCGACGAGAGAGAGGCAGAGATAAA	1440
Db	1381	AGAGAGGAAAGGATGAGGGAGAAAGAGAAAGAGGCGACGAGAGAGAGGCAGAGATAAA	1440
Qy	1441	GGAGGAGAGAAAAGGGCAGAGAGAGAGGCGAGAGATAAAGGAGGAAGAGAAAAGGGC	1500
Db	1441	GGAGGAGAGAAAAGGGCAGAGAGAGAGGCGAGAGATAAAGGAGGAAGAGAAAAGGGC	1500
Qy	1501	AGCAGAGAGAGGCGCAGAGATATAAGAGGAAAGAGAAAGGCGCAGAGAGAGAGGCAAG	1560
Db	1501	AGCAGAGAGAGGCGCAGAGATATAAGAGGAAAGAGAAAGGCGCAGAGAGAGAGGCAAG	1560
Qy	1561	GGCCAGGGCAGAGGAGAAAGAGAGAGGCTTAAGGTAGAAGAGCAGAAAAGCTTAAAGCAGCT	1620
Db	1561	GGCCAGGGCAGAGGAGAAAGAGAGAGGCTTAAGGTAGAAGAGCAGAAAAGCTTAAAGCAGCT	1620
Qy	1621	AGAAGAGAAAACATGCCATGCAAGAGACAAAGATATAAGGGGAAAGGTTAGACAGAA	1680
Db	1621	AGAAGAGAAAACATGCCATGCAAGAGACAAAGATATAAGGGGAAAGGTTAGACAGAA	1680
Qy	1681	AATAGAAGGAAATGGGTAAATCGAAAAGAAAGACAAAGAGATAAACTTACAGACAGCTGT	1740
Db	1681	AATAGAAGGAAATGGGTAAATCGAAAAGAAAGACAAAGAGATAAACTTACAGACAGCTGT	1740
Qy	1741	CCTAAAGAAAACAGGAGAGAGAGGGAACCTAAAGTGCACACTTAAAGAGAAAGGCTCCA	1800
Db	1741	CCTAAAGAAAACAGGAGAGAGAGGGAACCTAAAGTGCACACTTAAAGAGAAAGGCTCCA	1800
Qy	1801	AGAAGACAAGCCTACCTTCAAAAAGAAAGAGATCAAAAGATGAAAAAGATTTAAAAGGACAA	1860
Db	1801	AGAAGACAAGCCTACCTTCAAAAAGAAAGAGATCAAAAGATGAAAAAGATTTAAAAGGACAA	1860
Qy	1861	AGAACCCAAAGAAAGATTAAAGCTTCATGGATCGAAAGAGGATTTACAGAGCTTAA	1920
Db	1861	AGAACCCAAAGAAAGATTAAAGCTTCATGGATCGAAAGAGGATTTACAGAGCTTAA	1920
Qy	1921	GTGCGAGAAATCGAGAAATTCATGACCCCAAACTTTAAACATACTGAGAAATACTTTCAGCCG	1980
Db	1921	GTGCGAGAAATCGAGAAATTCATGACCCCAAACTTTAAACATACTGAGAAATACTTTCAGCCG	1980
Qy	1981	CCCTGGAGGGAGGCCAGCGTGAACACCAAGGAGCTGAGGGCGCCCCCAGGTGGAAGC	2040
Db	1981	CCCTGGAGGGAGGGCCAGCGTGAACACCAAGGAGCTGAGGGCGCCCCCAGGTGGAAGC	2040
Qy	2041	CGGCAAAAAGGCTGAGAGGCTTCGTTCGTGCGGGGAGAGCCGAGAGCGAAGAGTTTCGA	2100
Db	2041	CGGCAAAAAGGCTGAGAGGCTTCGTTCGTGCGGGGAGAGCCGAGAGCGAAGAGTTTCGA	2100
Qy	2101	GAAAGCTCAACAGAACGAGAGAGGAGGCGCTTTGGAGCTGGAGGAACCTCAAGAAAGAG	2160
Db	2101	GAAAGCTCAACAGAACGAGAGAGGAGGCGCTTTGGAGCTGGAGGAACCTCAAGAAAGAG	2160
Qy	2161	GGAGGAGAGAAAGGAGGTCTTGGAGGAGGAAGAGCAGAGGAGGAAGCAGGAGGAAGCCGA	2220
Db	2161	GGAGGAGAGAAAGGAGGTCTTGGAGGAGGAAGAGCAGAGGAGGAAGCAGGAGGAAGCCGA	2220
Qy	2221	TGCAAACTTCAGAGAGAGAGAGAGAGGCTTAAAGGAAGAGATTGAAAGGGCGAAG	2280
Db	2221	TGCAAACTTCAGAGAGAGAGAGAGAGGCTTAAAGGAAGAGATTGAAAGGGCGAAG	2280
Qy	2281	AGCAGAAAGCTGCTTGAGAAAAGCCCAAGAGATGCCAAGATGCTTGTTCAGATTCACAGAA	2340

2281 AGCAGAACTGCTGTGAGAAAGCGCAGAGAGATGCCAGAGATGGCTGTGCAGATGACAGAA 2340
2341 ACCATTCAAGTGTCTTCACTCTCTAAAGGTTTCATCTCTCAAGATAGAGAGAGCAGCAATT 2400
2341 ACCATTCAAGTGTCTTCACTCTCTAAAGGTTTCATCTCTCAAGATAGAGAGAGCAGCAATT 2400
2401 TTTGAAATGAGTGTGACAGAAAGCGTGGTGTCAATCGACCAATCAAGCAGCAATAGT 2460
2401 TTTGAAATGAGTGTGACAGAAAGCGTGGTGTCAATCGACCAATCAAGCAGCAATAGT 2460
2461 CTCGAAGATTGACAGCAGACTGGAGCAGTATACAGTGCATTTGAGGGAAACAAAAGCGC 2520
2461 CTCGAAGATTGACAGCAGACTGGAGCAGTATACAGTGCATTTGAGGGAAACAAAAGCGC 2520
2521 AAAACCTCAAAAGCGCGAGCGCTCGGATCTCTCTGTTCTCTGCTGAGGTGTACGCAACAT 2580
2521 AAAACCTCAAAAGCGCGAGCGCTCGGATCTCTCTGTTCTCTGCTGAGGTGTACGCAACAT 2580
2581 CAAGAGTATGTCGGAGAAAGGGAATGTTTTCATCCCCCACTGCAGCAGCAGCAACAAA 2640
2581 CAAGAGTATGTCGGAGAAAGGGAATGTTTTCATCCCCCACTGCAGCAGCAGCAACAAA 2640
2641 TAAGGAAACTGCTGCTGAAAGTAGGGTTTCTAGCGCATCAATGAATGGCTAACTAA 2700
2641 TAAGGAAACTGCTGCTGAAAGTAGGGTTTCTAGCGCATCAATGAATGGCTAACTAA 2700
2701 AACCCAGATGGAACAAGTCACTCTCTCCCAACCTTCTGACTTGAGACAGGAGACGT 2760
2701 AACCCAGATGGAACAAGTCACTCTCTCCCAACCTTCTGACTTGAGACAGGAGACGT 2760
2761 ATCCAGCAGCGGAACCTCTCGGAAAGCAATCTGTGATGAAGTCACTTCCCCCACTAA 2820
2761 ATCCAGCAGCGGAACCTCTCGGAAAGCAATCTGTGATGAAGTCACTTCCCCCACTAA 2820
2821 GGTGTGAGACGTTCCAGAAAGAACCCAGCTCAAGCAGCAGCAGCAGCTCAGTTGAGA 2880
2821 GGTGTGAGACGTTCCAGAAAGAACCCAGCTCAAGCAGCAGCAGCAGCTCAGTTGAGA 2880
2881 GGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTTACTAAATGGGTTCATTATCTTTT 2940
2881 GGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTTACTAAATGGGTTCATTATCTTTT 2940
2941 ATTTTCAATATCCAGTAAACCCATGATATATATCACTATATTTAATATCAAGTCTA 3000
2941 ATTTTCAATATCCAGTAAACCCATGATATATATCACTATATTTAATATCAAGTCTA 3000
3001 GAGATGTTTCATGTAAAGTACTGCTTTCACAGGAGCGCTGTTTCTAAAGAAACCCCATG 3060
3001 GAGATGTTTCATGTAAAGTACTGCTTTCACAGGAGCGCTGTTTCTAAAGAAACCCCATG 3060
3061 CTGTGAAATAGAGACTTTTCTACTGATCATCAATACTCTGTATCTGAGCAGTATACCAA 3120
3061 CTGTGAAATAGAGACTTTTCTACTGATCATCAATACTCTGTATCTGAGCAGTATACCAA 3120
3121 CCACATCTGAAGTCAACAGAAATCCAAATTTAAATGCTCGGAAATGTTGTCAGTAT 3180
3121 CCACATCTGAAGTCAACAGAAATCCAAATTTAAATGCTCGGAAATGTTGTCAGTAT 3180
3181 CTAGAAAAATGAACCGTAGTTTGTGTTTAAATACAGAAATCATGTTGTTCTGCAC 3240
3181 CTAGAAAAATGAACCGTAGTTTGTGTTTAAATACAGAAATCATGTTGTTCTGCAC 3240
3241 TTTATATTAAGCATGGAAGAAATTTCTAGTAGGCAATTTGACACTTTTGAAGTA 3300
3241 TTTATATTAAGCATGGAAGAAATTTCTAGTAGGCAATTTGACACTTTTGAAGTA 3300
3301 ACCATTTTCAGATTGAAATCTCGGATAATGTTGCTTTTAAAAAAGAAAAAGATGC 3360
3301 ACCATTTTCAGATTGAAATCTCGGATAATGTTGCTTTTAAAAAAGAAAAAGATGC 3360
3361 GGGCAGATTATTCCTTTGGTGGGTAAATTTACTTGCACTGGCGCGGTACAGTCGTG 3420
3361 GGGCAGATTATTCCTTTGGTGGGTAAATTTACTTGCACTGGCGCGGTACAGTCGTG 3420

3421 CTGGGAAACTTGGCGTCCCAATATTCGCATGTCGGAATCCCTTCGACGCTGGTATGCGAA 3480
3421 CTGGGAAACTTGGCGTCCCAATATTCGCATGTCGGAATCCCTTCGACGCTGGTATGCGAA 3480
3481 GGCACCATGCTCCAAATGCGGTGATGGATGGGGCTGTGCTTACGCGGTTTGGGT 3540
3481 GGCACCATGCTCCAAATGCGGTGATGGATGGGGCTGTGCTTACGCGGTTTGGGT 3540
3541 TCCGACCGAAATTCAGCACCGGGCTTCTCTTTTGGGGGGGGCCCTTAAAGGGCTTG 3600
3541 TCCGACCGAAATTCAGCACCGGGCTTCTCTTTTGGGGGGGGCCCTTAAAGGGCTTG 3600
3601 GGGTTTTTCCGCCGAAATTTGGCTTGTGCCGAAATCCCTTTGGGGTTACGCTGCACACAC 3660
3601 GGGTTTTTCCGCCGAAATTTGGCTTGTGCCGAAATCCCTTTGGGGTTACGCTGCACACAC 3660
3661 GTTTTGTCCGGAGAGGAAAGCAAGGGCGGTTACCGAGAAC 3702
3661 GTTTTGTCCGGAGAGGAAAGCAAGGGCGGTTACCGAGAAC 3702

RESULT 2

US-10-062-674-1815
; Sequence 1815, Application US/10062674
; Publication No. US2004000559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/625,102
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PESTL Program
; SEQ ID NO 1815
; LENGTH: 3500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US2004000559A1 256147.56
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (3500)
; OTHER INFORMATION: a, t, c, g, or other
US-10-062-674-1815

Query Match 86.5%; Score 3201.2; DB 15; Length 3500;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 3206; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 126 GGGATATGTGTTCACTTAGCATGACTTCTGGAGGGGCCAAGAAAGGGCGGTCTGGAGT 185
DB 1 GGGATATGTGTTCACTTAGCATGACTTCTGGAGGGGCCAAGAAAGGGCGGTCTGGAGT 60
QY 186 TTTATTGATAGACAGTGTGTTATTCGGCTGCTGCCCTGCCCTGCTCTCTCTGCTG 245
DB 61 TTTATTGATAGACAGTGTGTTATTCGGCTGCTGCCCTGCCCTGCTCTCTCTGCTG 120
QY 246 TGTGCTCTGCTTTAAAGAAATCAGTCTCTTTCCGACTTAGTCTCGGGAAGAGTTT 305
DB 121 TGTGCTCTGCTTTAAAGAAATCAGTCTCTTTCCGACTTAGTCTCGGGAAGAGTTT 180
QY 306 CAGACTCAAGGTATCATTTGGAACATTTCAAGATCATCAATCAAAATCCACAGGATG 365
DB 181 CAGACTCAAGGTATCATTTGGAACATTTCAAGATCATCAATCAAAATCCACAGGATG 240
QY 366 GTGACCAACCAAGGCTCAGACATCTGATTTGCTGACCTGTCCAGACATCATCTGCTC 425
DB 241 GTGACCAACCAAGGCTCAGACATCTGATTTGCTGACCTGTCCAGACATCATCTGCTC 300

QY 426 CCTGAACCTGAATCACACCATGGATGATTTTGGCGTCCAGAGAACTTGAAGGCCAAA 485
 Db 301 CCTGAACTCTGAATCACACCATGGATGATTTTGGCGTCCAGAGAACTTGAAGGCCAAA 360
 QY 486 AGAGGGAGGAGATCGGACTCGAGCAGAGAAAGATCGCTTACAGAGGAATGACGATG 545
 Db 361 AGAGGGAGGAGATCGGACTCGAGCAGAGAAAGATTCCTTACAGAGGAATGACGATG 420
 QY 546 AAGAGGAGGAGCCGGGAAACGGCCGCGCGAGCCCGACAGAAACGGCTCGGCGCAAGC 605
 Db 421 AAGAGGAGGAGCCGGGAAACGGCCGCGCGAGCCCGACAGAAACGGCTCGGCGCAAGC 480
 QY 606 AGAGGAAGATCTTGGGACAGGTGACCGCAGCGTGGAGTGAATGCCCAGAACAGTG 665
 Db 481 AGAGGAAGATCTTGGGACAGGTGACCGCAGCGTGGAGTGAATGCCCAGAACAGTG 540
 QY 666 TGCCTGACGAGAGGCCCAAGACAACCCACCAACACTCAAGTGAAGGGGATGATGAGG 725
 Db 541 TGCCTGACGAGAGGCCCAAGACAACCCACCAACACTCAAGTGAAGGGGATGATGAGG 600
 QY 726 CCGCATTCCTGGAGGCGCTGGCTCGGCTGAGGAAGACGCCAAAACGGCTTCAGGAGG 785
 Db 601 CCGCATTCCTGGAGGCGCTGGCTCGGCTGAGGAAGACGCCAAAACGGCTTCAGGAGG 660
 QY 786 CTCTGGAGCGGAGAGGATTCGACCCCAACAATAACAGATGCAAGTCTGCTCTCCAA 845
 Db 661 CTCTGGAGCGGAGAGGATTCGACCCCAACAATAACAGATGCAAGTCTGCTCTCCAA 720
 QY 846 GCAGAGAAATGCAAAATGACACAGCGAAATATGAACCTACCGAGAGGAAGAAAAGTG 905
 Db 721 GCAGAGAAATGCAAAATGACACAGCGAAATATGAACCTACCGAGAGGAAGAAAAGTG 780
 QY 906 AAAGTCGCCAAGAAAGATACGAGATAGGAAACAGAAACAGTCAACCAAGTCTTACCAGA 965
 Db 781 AAAGTCGCCAAGAAAGATACGAGATAGGAAACAGAAACAGTCAACCAAGTCTTACCAGA 840
 QY 966 AGAATGATTTGGAGGATGCTGAAGAAAACAAGAAAGAGACAGAAAGAGAGGAGG 1025
 Db 841 AGAATGATTTGGAGGATGCTGAAGAAAACAAGAAAGAGACAGAAAGAGAGGAGG 900
 QY 1026 AAGAGGAGAGCCAAAGCCAGGAGGATTCGAGAAATCAGGTAGAGGTGATGCTGGAAG 1085
 Db 901 AAGAGGACAAAGCCAAAGCCAGGAGGATTCGAGAAATCAGGTAGAGGTGATGCTGGAAG 960
 QY 1086 AGAAACCACTGAAGCCAGGAGGAAACAGTGTGAATGTCTTAAAAAATGGGCAGATCA 1145
 Db 961 AGAAACCACTGAAGCCAGGAGGAAACAGTGTGAATGTCTTAAAAAATGGGCAGATCA 1020
 QY 1146 GTTCAGAGAGCTTAACAGAGGAGGAGGAGGAAACAGGTTCAGATGAGATTTCCCATC 1205
 Db 1021 GTTCAGAGAGCTTAACAGAGGAGGAGGAGGAAACAGGTTCAGATGAGATTTCCCATC 1080
 QY 1206 ATGAAAGATGGAAGAGGAGAGACAAGGAAAGAGCTGAGGAGAGAGGCGCAAGTTGGAAG 1265
 Db 1081 ATGAAAGATGGAAGAGGAGAGACAAGGAAAGAGCTGAGGAGAGAGGCGCAAGTTGGAAG 1140
 QY 1266 CAGAGAAAGAGAAAGATTAAGCCGAGCAAGCAAAAGATAGCAGATGAACGACCA 1325
 Db 1141 CAGAGAGCAGAGAGAAATTAAGCCGAGCAAGCAAAAGATAGCAGATGAACGAGCA 1200
 QY 1326 GAATTCAGCAGAAAGAAAGAGCTGCCCAAGAAAGAGAGAGAGAGGAGGAGAGAGA 1385
 Db 1201 GAATTCAGCAGAAAGAAAGAGCTGCCCAAGAAAGAGAGAGAGAGGAGGAGAGAGA 1260
 QY 1386 GGGAAAGATGAGGAGGAGAAAGAAAGCCGAGCAGAGAGAGGAGCAGAGATTAAGGAGG 1445
 Db 1261 GGGAAAGATGAGGAGGAGAAAGAAAGCCGAGCAGAGAGAGGAGCAGAGATTAAGGAGG 1320
 QY 1446 AAGAGAAAGGAGCAGAGAGGAGGAGGAGGATTAAGGAGGAGAGAGAAAGGAGGAGCAG 1505
 Db 1321 AAGAGAAAGGAGCAGAGAGGAGGAGGAGGATTAAGGAGGAGAGAGAAAGGAGGAGCAG 1380
 QY 1506 AAGAGAGGAGAGGATTAAGAGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1565

Db 1381 AGGAGAGGAGAGGATTAAGAGAGAGAAAGGGCGAGAGGAGGAGGAGGAGGAGGAGG 1440
 QY 1566 GGGCAGAGGAGGAGAGAGAGGCTAAGGTAGAGAGCAGAGAAACGTAACAAGAGCTAGAAG 1625
 Db 1441 GGGCAGAGGAGGAGAGAGAGGCTAAGGTAGAGAGCAGAGAAACGTAACAAGAGCTAGAAG 1500
 QY 1626 AGAATAAATCATGCCATGCCAGAGCAAAAGATTAAGAGGAGGAGGAGGAGGAGGAGGAG 1685
 Db 1501 AGAATAAATCATGCCATGCCAGAGCAAAAGATTAAGAGGAGGAGGAGGAGGAGGAGGAG 1560
 QY 1686 AAGGGAATGGGTAAATGAAAGAAAGCACAAGAAAGATAAACTTCAGACAGCTGTCTAA 1745
 Db 1561 AAGGGAATGGGTAAATGAAAGAAAGCACAAGAAAGATAAACTTCAGACAGCTGTCTAA 1620
 QY 1746 AGAAGAGGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1805
 Db 1621 AGAAGAGGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
 QY 1806 ACAAGCCTTACCTTCAAAAAGAGAGATCAAAAGATGAAAGATTAAGAGGAGGAGGAGGAG 1865
 Db 1681 ACAAGCCTTACCTTCAAAAAGAGAGATCAAAAGATGAAAGATTAAGAGGAGGAGGAGGAG 1740
 QY 1866 CCAAGAGAGAGGATTAAGAGCTTCATGATCGAAAGAGGAGGATTAAGAGGAGGAGGAGGAG 1925
 Db 1741 CCAAGAGAGAGGATTAAGAGCTTCATGATCGAAAGAGGAGGATTAAGAGGAGGAGGAGGAG 1800
 QY 1926 AGAATGGAGATTCATGACCCCAACAACTTAAACATACATCTGAGGAGGAGGAGGAGGAG 1985
 Db 1801 AGAATGGAGATTCATGACCCCAACAACTTAAACATACATCTGAGGAGGAGGAGGAGGAG 1860
 QY 1986 GAG 2045
 Db 1861 GAG 1920
 QY 2046 AAAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2105
 Db 1921 AAAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1980
 QY 2106 TCAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2165
 Db 1981 TCAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2040
 QY 2166 AGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2225
 Db 2041 AGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
 QY 2226 AACTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2285
 Db 2101 AACTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
 QY 2286 AAGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2345
 Db 2161 AAGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
 QY 2346 TCAAGTGTTCCT 2405
 Db 2221 TCAAGTGTTCCT 2280
 QY 2406 ATAAGTCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2465
 Db 2281 ATAAGTCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
 QY 2466 AGATTGACAGCAGACCTGGAGCAGTATACCAAGTGCATTTGAGGAGGAGGAGGAGGAGGAGGAGGAG 2525
 Db 2341 AGATTGACAGCAGACCTGGAGCAGTATACCAAGTGCATTTGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
 QY 2526 CTACAAAGCCGAGCAGCTCGGATCT 2585
 Db 2401 CTACAAAGCCGAGCAGCTCGGATCT 2460
 QY 2586 GTATGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2645

QY	1837	AGATGAAAAGATTAAAAAGGACAAAGAACCCAAAGAAAGTTAAAGAGCTTTCATGGATCG	1896
DB	1621	AGATGAAAAGATTAAAAAGGACAAAGAACCCAAAGAAAGTTAAAGAGCTTTCATGGATCG	1680
QY	1997	AAAGAAAGGATTTACAGAAAGTTAAGTCGACGAATGGAGAAATTCATGACCCACAAACTTAA	1956
DB	1681	AAAGAAAGGATTTACAGAAAGTTAAGTCGACGAATGGAGAAATTCATGACCCACAAACTTAA	1740
QY	1957	ACATACTCAGAAATACCTTTTCAGCGCCCTCGAGGGAGGCGCAGCGTGGACACCAAGGAGGC	2016
DB	1741	ACATACTCAGAAATACCTTTTCAGCGCCCTCGAGGGAGGCGCAGCGTGGACACCAAGGAGGC	1800
QY	2017	TGAGGGCCCCCCCCAGGTGGAGCGGCAAAAGSGCTGGAGGAGCTTCGTCTGCTCGCGG	2076
DB	1801	TGAGGGCCCCCCCCAGGTGGAGCGGCAAAAGSGCTGGAGGAGCTTCGTCTGCTCGCGG	1860
QY	2077	GGAGACCGAGACGCGAAGAGTTTCGAGAAAGCTCAAAACAGAACGACGAGAGGCGGCTTTGGGA	2136
DB	1861	GGAGACCGAGACGCGAAGAGTTTCGAGAAAGCTCAAAACAGAACGACGAGAGGCGGCTTTGGGA	1920
QY	2137	GCTGGAGGAACTCAAGAAAAAGAGGAGGAGAGAAAGGAAGGTCTCTGGAGGAGGAAGAGCA	2196
DB	1921	GCTGGAGGAACTCAAGAAAAAGAGGAGGAGAGAAAGGAAGGTCTCTGGAGGAGGAAGAGCA	1980
QY	2197	GAGGAGGAAGCAGGAGCGAAAGCCGATCGAANAACCTCAGAGAGGAGGAAGAGAGAGGAGGCT	2256
DB	1991	GAGGAGGAAGCAGGAGGAGAAAGCCGATCGAANAACCTCAGAGAGGAGGAAGAGAGAGGCT	2040
QY	2257	AAAGGAAGAGATTGAAAGGCGAAAGAGACGAAGTCTGTGAGAAAACGCCAGAAAGATGCCAGA	2316
DB	2041	AAAGGAAGAGATTGAAAGGCGAAAGAGACGAAGTCTGTGAGAAAACGCCAGAAAGATGCCAGA	2100
QY	2317	AGATGCTGTTCAGATGACAAAGAAAACCATTCAGTGTTTCACTCTAAAGGTTTCATCTCT	2376
DB	2101	AGATGCTGTTCAGATGACAAAGAAAACCATTCAGTGTTTCACTCTCTAAAGGTTTCATCTCT	2160
QY	2377	CAAGATAGAAAGACGAGCAGAAATTTTGAATAGTCTGTGCAGAAAAGCAGTGGTGTCAA	2436
DB	2161	CAAGATAGAAAGACGAGCAGAAATTTTGAATAGTCTGTGCAGAAAAGCAGTGGTGTCAA	2220
QY	2437	ATTCGACCCATCAAGCAGCAATAGTCTCCAAGATTGACAGCAGACTGGAGCAGTATACCAG	2496
DB	2221	ATTCGACCCATCAAGCAGCAATAGTCTCCAAGATTGACAGCAGACTGGAGCAGTATACCAG	2280
QY	2497	TGCAATTTAGGGGAACAAAAGCCGAAAACCTTACAAAAGCCGACGCTCGGATCTTCTGT	2556
DB	2281	TGCAATTTAGGGGAACAAAAGCCGAAAACCTTACAAAAGCCGACGCTCGGATCTTCTGT	2340
QY	2557	TCCTGCTGAAGGTTGACGCAACATCAAGAGPATGTGGAGAAAAGGGAATGTGTTTTTCATC	2616
DB	2341	TCCTGCTGAAGGTTGACGCAACATCAAGAGPATGTGGAGAAAAGGGAATGTGTTTTTCATC	2400
QY	2617	CCCCACTGCAGCAGGACACCCAAATTAAGAAAACCTGCTGGCTTGAAGGTAGGGGTTTCTAG	2676
DB	2401	CCCCACTGCAGCAGGACACCCAAATTAAGAAAACCTGCTGGCTTGAAGGTAGGGGTTTCTAG	2460
QY	2677	CGGCATCAATGAATGGCTTAATAAACCCAGATGGAACCAAGTCACCTGCTCCCAAAACC	2736
DB	2461	CGGCATCAATGAATGGCTTAATAAACCCAGATGGAACCAAGTCACCTGCTCCCAAAACC	2520
QY	2737	TTCTGACTTCGAGACGAGGAGACGTATCCAGCAAGCGGAAACCTCTCGGAAAAAGCAATCTGT	2796
DB	2521	TTCTGACTTCGAGACGAGGAGACGTATCCAGCAAGCGGAAACCTCTCGGAAAAAGCAATCTGT	2580
QY	2797	GGATAGGTCACTTCCCCCACTTAAGTTTGGACAGAGTTCCAGAAAGAACCCAAAGCTCAAAG	2856
DB	2581	GGATAGGTCACTTCCCCCACTTAAGTTTGGACAGAGTTCCAGAAAGAACCCAAAGCTCAAAG	2640
QY	2857	ACGCAGGACGAGCTCAGTGTAGAGGGCTAAATTCGCTCTGTTTGTATTTAGTTGATTT	2916
DB	2641	ACGCAGGACGAGCTCAGTGTAGAGGGCTAAATTCGCTCTGTTTGTATTTAGTTGATTT	2700

Qy	2917	ACTAAATGGGTTCAATATCTTTTAAATTTTCAATATCCCGAGTAAACCCGATATATTATC	2976
Db	2701	ACTAAATGGGTTCAATATCTTTTAAATTTTCAATATCCCGAGTAAACCCGATATATTATC	2760
Qy	2977	ACTATATTTTAATAATCAGAGTCTAGAGATGTTTCATGGTAAAGTACTGCCTTTGCACAGG	3036
Db	2761	ACTATATTTTAATAATCAGAGTCTAGAGATGTTTCATGGTAAAGTACTGCCTTTGCACAGG	2820
Qy	3037	AGCCTGTTTCTAAAGAAAACCCATGCTGTGAAATPAGAGACTTTTCTACTGATCATCATAAAC	3096
Db	2821	ATCCTGTTTCTAAAGAAAACCCATGCTGTGAAATPAGAGACTTTTCTACTGATCATCATAAAC	2880
Qy	3097	TCTGTATCTGACGATGATACCAACACCATCTGAAAGTCAACAGAGAAGTCCAAAGTTTAAAA	3156
Db	2881	TCTGTATCTGACGATGATACCAACACCATCTGAAAGTCAACAGAGAAGTCCAAAGTTTAAAA	2940
Qy	3157	TTGCCCTGGGGAATGTGTGCAGTATCTAGAAAAATGAAACCGTAGTTTGTTTTTTTTAAAT	3216
Db	2941	TTG-CTGGGAATGTGTGCAGTATCTAGAAAAATGAAACCGTAGTTTGTTTTTTTTAAAT	2999
Qy	3217	ACAGAAGTCATGTTGTTTCTGCACTTTATATATAAAGCATGGAAGAAATTTATCTTAGTAGG	3276
Db	3000	ACAGAAGTCATGTTGTTTCTGCACTTTATATATAAAGCATGGAAGAAATTTATCTTAGTAGG	3059
Qy	3277	CAATTTGTACACTTTTGAAGTAAACCCATTTCCAGATTTGAAATCTCGGATAATGGTTG	3336
Db	3060	CAATTTGTACACTTTTGAAGTAAACCCATTTCCAGATTTGAAATCTCGCAATAATGGTTG	3119
Qy	3337	CTTTAAAAAAGAAAAAGATCGGGCAGTATTTCCTTT	3378
Db	3120	CTTTTAAAAAAGAAAAAGATGTAAGTCTGTTAAAGGTAATCTTTT	3161

RESULT 5

```

US-10-062-674-1817/C
; Sequence 1817, Application US/10062674
; Publication No. US2004000559A1
;
; GENERAL INFORMATION:
; APPLICANT: LOIING, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND P
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/62,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/525,102
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PERL Program
; SEQ ID NO 1817
; LENGTH: 4566
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US2004000559A1 256147.586
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (4566)
; OTHER INFORMATION: a, t, c, g, or other
; US-10-062-674-1817

```

Query Match	Score	DB	Length
Best Local Similarity	63.6%	2353	4566
Matches 2616; Conservative	52.0%	Pred. No. 0;	
Matches 2616; Conservative	0	Mismatches	11
		Indels	217
		Gaps	4

QY	517	AATCGCCTTACCAGAGGAATGACGATGATGAAGAGGAGCGACGCCGGGAAACGGCGCCGCCG	576
DB	4474	AATCGCCTTACCAGAGGAATGACGATGATGAAGAGGAGCGACGCCGGGAAACGGCGCCGCCG	4418
QY	577	AGCCCGCAGAGAAACGGCTGCGCGCAGAGCAGAGGAGAAATCTCTTGGGACAGGTCACCGA	636
DB	4415	AGCCCGCAGAGAAACGGCTGCGCGCAGAGCAGAGGAGAAATCTCTTGGGACAGGTCACCGN	435
QY	637	CCAGTGTGAGGTGTAATCGCCAGAAACAGTGTGCCCTGACGAGGAGGCCAAGAACAACACCAAC	696

QY	2856	GAGCCAGGACGAGCTCAGTTGTAGAGGGTAAATTCGTCCTGTTTGTATTATTGTTGATT	2913
DB	2172	-----	2173
QY	2916	TACTAAATTGGGTTTCATTATCTTTTATTTTCAATATCCCAAGTAAACCCATGTATATTAT	2975
DB	2172	-----	2173
QY	2976	CACATATATTTAATAATCACAGCTCTAGAGATGTTCTAGTGTAAAAAGTACTGCTTTTGCCAG	3035
DB	2172	-----	2171
QY	3036	GAGCGTGTCTTAAAGAAACCCGTCGTGCAATAGAGACTTTTCTACTGATCATCATATA	3095
DB	2170	GAGCGTGTCTTAAAGAAACCCGTCGTGCAATAGAGACTTTTCTACTGATCATCATATA	2111
QY	3096	CTCTGTATCTGACGAGTGATACCAACACATCTGAAGTCAACAGAGATCCAAAGTTTAAA	3155
DB	2110	CTCTGTATCTGACGAGTGATACCAACACATCTGAAGTCAACAGAGATCCAAAGTTTAAA	2051
QY	3156	ATTGCCCTGCCGAATGCTGCGAGTATCTAGAAAAATGAACCGTAGTTTTCGTTTTTTAAA	3215
DB	2050	ATTGCCCTGCCGAATGCTGCGAGTATCTAGAAAAATGAACCGTAGTTTTCGTTTTTTAAA	1991
QY	3216	TACAGAAAGTCATGTTGTTTCTGCACTTTATAATAAAGCATGGAAGAAATATCTTAGTAG	3275
DB	1990	TACAGAAAGTCATGTTGTTTCTGCACTTTATAATAAAGCATGGAAGAAATATCTTAGTAG	1931
QY	3276	GCAATTCGAACACATTTTGAAGTAAACCCATTTCAGATTTTGAATACTGGCGATAATGGTT	3335
DB	1930	GCAATTCGAACACATTTTGAAGTAAACCCATTTCAGATTTTGAATACTGGCGATAATGGTT	1871
QY	3336	GCCTTTAAAAAAGAAAGATG 3359	
DB	1870	GTCTTTAAAAAAGAAAGATG 1847	
RESULT 6			
US-09-974-298-36			
; Sequence 36, Application US/09974298			
; Patent No. US20020156263A1			
; GENERAL INFORMATION:			
; APPLICANT: Chen, Hwei-Mei			
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER			
; FILE REFERENCE: PA-0037 P			
; CURRENT APPLICATION NUMBER: US/09/974,298			
; CURRENT FILING DATE: 2001-10-04			
; PRIOR APPLICATION NUMBER: 60/238,331			
; PRIOR FILING DATE: 2000-05-10			
; NUMBER OF SEQ ID NOS: 194			
; SOFTWARE: PERL Program			
; SEQ ID NO 36			
; LENGTH: 4547			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc_feature			
; OTHER INFORMATION: Incyte ID No. US20020156263A1 022404.25			
US-09-974-298-36			
Query Match 48.4%; Score 1793.6; DB 9; Length 4547;			
Best Local Similarity 76.9%; Pred. No. 0;			
Matches 2585; Conservative 0; Mismatches 9; Indels 766; Gaps 2;			
QY	1	GCCTAGGGAATGACAGGCATCTCCACAGCAGGCTGCATCCACCTT-GGCTGGGGTGTCG	59
DB	84	GCCTAGGGAATGACAGGCATCTCCACAGCAGGCTGCATCCACCTTGGGCTGGGGTGTCG	143
QY	60	TCATTGGCTGCCATTAGAAAAACGACAGGCAATGTCATCCACCGCTCCCGACTGTAA	119
DB	144	TCATTGGCTGCCATTAGAAAAACGACAGGCAATGTCATCCACCGCTCCCGACTGTAA	203

Db 1149 ----- 1148
 Qy 1260 TGAAGCAGAAAGAGAGAAAGAAATTAAGCCGAGCAGACAAAGAGATAGCAGATCAAC 1319
 Db 1149 ----- 1148
 Qy 1320 GAGCAAGAATTGAAGCAGAAAGAAAGACAGCTGCCAAGAAAGAGAAAGGAGAGAGCAG 1379
 Db 1149 ----- 1148
 Qy 1380 AAGAGAGGAAAGGATGAGGGAGAAAGAGAAAGGAGCAGAGGAGGAGGAGGATAA 1439
 Db 1149 ----- 1148
 Qy 1440 AGGAGAGAGAAAGGCGCAGCAGAGAGGAGGAGAGGATTAAGGAGGAGAGAAAGGG 1499
 Db 1149 ----- 1148
 Qy 1500 CAGCAGAGGAGGAGCAGAGGATTAAGAGAGGAGAAAGGAGCAGAGAGGAGGCAAA 1559
 Db 1149 ----- 1148
 Qy 1560 GGGCCAGGCGAGAGGAGAGAGAGAGGCTTAAGTAGAGAGCAGAGACGTAACAGCAGC 1619
 Db 1149 ----- 1148
 Qy 1620 TAGAAGAGAAAAACATGCCATGCAAGAGACAAAGATAAAAAAGGGAAGAGGTAGACAGA 1679
 Db 1149 ----- 1148
 Qy 1680 AAATAGAGGAAATGGGTAAATGAAAAAGAAACAAAGAGATAAATCTCAGACAGCTG 1739
 Db 1149 ----- 1148
 Qy 1740 TCCTAAGAAACAGGAGAGAGAGAGGAACTTAAGTGCAAGCTTAAGAGAAAGAGCTCC 1799
 Db 1149 ----- 1148
 Qy 1800 AAGAGACAAAGCTTACCTTCAAAAAAGAGAGATCAAGAGTAAAGATTAAAAAGGACA 1859
 Db 1149 ----- 1148
 Qy 1860 AAGAACCCAAAGAGAGATTAGAGCTTCATGATCGAAGAGAGGATTTACAGAGTTA 1919
 Db 1179 AAGAACCCAAAGAGAGATTAGAGCTTCATGATCGAAGAGAGGATTTACAGAGTTA 1238
 Qy 1920 AGTCGAGAAATGAGAAATTCATGACCCACAAACTTAAACATACCTGAGAAATCTTTCAGCC 1979
 Db 1239 AGTCGAGAAATGAGAAATTCATGACCCACAAACTTAAACATACCTGAGAAATCTTTCAGCC 1298
 Qy 1980 GGCCTGAGAGGAGGCGCAGCTGGAACCAAGAGGCTGAGGCGCCGCCAGAGTGGAG 2039
 Db 1299 GGCCTGAGAGGAGGCGCAGCTGGAACCAAGAGGCTGAGGCGCCGCCAGAGTGGAG 1358
 Qy 2040 CCGGCAAAAGGCTGAGAGGCTTCGTCGTCGCGGGAGACGAGAGCAAGATTGCG 2099
 Db 1359 CCGGCAAAAGGCTGAGAGGCTTCGTCGTCGCGGGAGACGAGAGCAAGATTGCG 1418
 Qy 2100 AGAAGCTCAAAACAGAGCAGCAGAGGCGGCTTTGAGCTGGAGAACTCAAGAAAAAGA 2159
 Db 1419 AGAAGCTCAAAACAGAGCAGCAGAGGCGGCTTTGAGCTGGAGAACTCAAGAAAAAGA 1478
 Qy 2160 GGGAGAGAGAGAGGCTTCGAGGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGCG 2219
 Db 1479 GGGAGAGAGAGAGGAGGCTTCGAGGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGCG 1538
 Qy 2220 ATCGAAAACTCAGAGAGGAGAGAGAGAGGAGGCTTAAAGAGAGAGATTGAAAGGCGAA 2279
 Db 1539 ATCGAAAACTCAGAGAGGAGAGAGAGAGAGGAGGCTTAAAGAGAGAGATTGAAAGGCGAA 1598
 Qy 2280 GAGCAAGAGCTGCTGAGAAAGCCAGAGAGATGCCAGAGAGTGGCTTGTACATGACAGA 2339

Db 1599 GAGCAGAGCTGCTGAGAAACGCCAGAGATGCCAGAGATGGCTTGTACATGACAAGA 1658
 Qy 2340 AACCAATCAAGTGTCTTCACTCTTAAGAGTTCATCTCTCAAGATAGAGAGCGCAGAGAT 2399
 Db 1659 AACCAATCAAGTGTCTTCACTCTTAAGAGTTCATCTCTCAAGATAGAGAGCGCAGAGAT 1718
 Qy 2400 TTTTGAATAAGTCTGTGAGAAAGCAGTGGTGTCAAAATCGACCATCAAGCAGCAATAG 2459
 Db 1719 TTTTGAATAAGTCTGTGAGAAAGCAGTGGTGTCAAAATCGACCATCAAGCAGCAATAG 1778
 Qy 2460 TCTCAAGATTGACAGCAGACTGGAGCAGATATACCAAGTGAATTGAGGGAAACAAAGCG 2519
 Db 1779 TCTCAAGATTGACAGCAGACTGGAGCAGATATACCAAGTGAATTGAGGGAAACAAAGCG 1838
 Qy 2520 CAAAACCTTACAAAGCGCGCAGCTCGATCTTCTCTGTCTCTGCTGAAGTGTACGCAACA 2579
 Db 1839 CAAAACCTTACAAAGCGCGCAGCTCGATCTTCTCTGTCTCTGCTGAAGTGTACGCAACA 1898
 Qy 2580 TCAAGAGTATGGGAGAAAGGAATGTGTTCATCTCCCACTGCGAGCAGCAGCAACAA 2639
 Db 1899 TCAAGAGTATGGGAGAAAGGAATGTGTTCATCTCCCACTGCGAGCAGCAGCAACAA 1958
 Qy 2640 ATAAGAACTCTGCTGTGAGAGTGGGTGTCTAGCCCATCAATGAATGGCTAACTA 2699
 Db 1959 ATAAGAACTCTGCTGTGAGAGTGGGTGTCTAGCCCATCAATGAATGGCTAACTA 2018
 Qy 2700 AAACCCCGAGATGGAACCAAGTCACTGCTGCCCAACCTTCTGACTTGAGACAGGAGAGC 2759
 Db 2019 AAACCCCGAGATGGAACCAAGTCACTGCTGCCCAACCTTCTGACTTGAGACAGGAGAGC 2078
 Qy 2760 TATCAGCAGAGGAGACCTCTGGGAAAGCAATCTGTGGATAGGTCACTTCCCCACTA 2819
 Db 2079 TATCAGCAGAGGAGACCTCTGGGAAAGCAATCTGTGGATAGGTCACTTCCCCACTA 2138
 Qy 2820 AGGTTTGACACAGTTCAGAAAGAACCCAGCTCAAGACGAGCAGAGCTCAGTTGTAG 2879
 Db 2139 AGGTTTGACACAGTTCAGAAAGAACCCAGCTCAAGACGAGCAGAGCTCAGTTGTAG 2198
 Qy 2880 AGGCTAATTCGCTCTGTTTGTATTTATGTTGATTTACTAAATGGTTCATTTCTTT 2939
 Db 2199 AGGCTAATTCGCTCTGTTTGTATTTATGTTGATTTACTAAATGGTTCATTTCTTT 2258
 Qy 2940 TATTTTCAATATCCAGTAAACCCATGATATTTACTCATATATTTAATAATCACAGTCT 2999
 Db 2259 TATTTTCAATATCCAGTAAACCCATGATATTTACTCATATATTTAATAATCACAGTCT 2318
 Qy 3000 AGAGATGTTTCATGTTAAAGTACTGCTTTGCAAGAGGCTGTTCTTAAAGAAACCCAT 3059
 Db 2319 AGAGATGTTTCATGTTAAAGTACTGCTTTGCAAGAGGCTGTTCTTAAAGAAACCCAT 2378
 Qy 3060 GCTGTGAATAGAGACTTTTCTACTGATCATCACTCTGTATCTGTATCTGAGCAGTATACCA 3119
 Db 2379 GCTGTGAATAGAGACTTTTCTACTGATCATCACTCTGTATCTGTATCTGAGCAGTATACCA 2438
 Qy 3120 ACCACATCTGAAGTCAACAGAGATCCAAGTTTAAATTTGCTCGGAATGTGTGAGTA 3179
 Db 2439 ACCACATCTGAAGTCAACAGAGATCCAAGTTTAAATTTGCTCGGAATGTGTGAGTA 2498
 Qy 3180 TCTAGAAAAATGAACGCTAGTTTTGTTTTTAAATACAGAGTCACTGTTGTTTCTGCA 3239
 Db 2499 TCTAGAAAAATGAACGCTAGTTTTGTTTTTAAATACAGAGTCACTGTTGTTTCTGCA 2558
 Qy 3240 CTTTATAATAAGCATGGAAGAAATTTATCTAGTAGGCAATTTGAAACACTTTTGTAAAGT 3299
 Db 2559 CTTTATAATAAGCATGGAAGAAATTTATCTAGTAGGCAATTTGAAACACTTTTGTAAAGT 2618
 Qy 3300 AACCCATTTTCAGATTTGAAATACCTGCGATTAATGGTTCGTTTTAAAGAAAAAGAAAGATG 3359
 Db 2619 AACCCATTTTCAGATTTGAAATACCTGCAATTAATGGTTCGTTTTTAAAGAAAAAGAAATG 2678

D	b	1149	- - - - -	1148
Q	y	1800	AAGAAGACAAGCCTACCTTCAA AAAAAGAGAGATCAAAAGATGAAAAGATTAAAAAGGACA	1859
D	b	1149	- - - - -	1148
Q	y	1860	AAGAACCCAAAGAGGAAGTTAAGAGCTTCATGGATCGAAAGAAAGGATTTACAGAGTTTA	1919
D	b	1179	AAGAACCCAAAGAGGAAGTTAAGAGCTTCATGGATCGAAAGAGGATTTACAGAGTTTA	1238
Q	y	1920	AGTCGCAGAATGGAGAAATCATGACCCACAAACTTAAACATACTGAGAAATACTTTTCAGCC	1979
D	b	1239	AGTCGCAGAATGGAGAAATCATGACCCACAAACTTAAACATACTGAGAAATACTTTTCAGCC	1298
Q	y	1980	GCCCTGGAGGAGGCCAGCGTGACACCAAGGAGGCTTGAGGGCGCCCCCNAGGTGGAG	2039
D	b	1299	GCCCTGGAGGAGGCCAGCGTGACACCAAGGAGGCTTGAGGGCGCCCCCNAGGTGGAG	1358
Q	y	2040	CGGCAAAAGGCTGGAGGAGCTTCGTGCTCGCGGGGAGACCAGAGCGAAGAGTTCG	2099
D	b	1359	CGGCAAAAGGCTGGAGGAGCTTCGTGCTCGCGGGGAGACCAGAGCGAAGAGTTCG	1418
Q	y	2100	AGAAGCTCAAAACAGAGACGAGAGCGCGCTTTGGAGCTGGAGGAATCTCAAGAAAAGA	2159
D	b	1419	AGAAGCTCAAAACAGAGACGAGAGCGCGCTTTGGAGCTGGAGGAATCTCAAGAAAAGA	1478
Q	y	2160	GGGAGGAGAGGAAGGTCCTCGAGGAGGAAGAGCAGAGAGGAAGCAGGAGGAAGCGC	2219
D	b	1479	GGGAGGAGAGGAAGGTCCTCGAGGAGGAAGAGCAGAGAGGAAGCAGGAGGAAGCGC	1538
Q	y	2220	ATCGAAAACCTCAGAGAGGAGGAAGAGAGGAGGCTTAAAGGAAGAGATTTGAAAGCGGAA	2279
D	b	1539	ATCGAAAACCTCAGAGAGGAGGAAGAGAGGAGGCTTAAAGGAAGAGATTTGAAAGCGGAA	1598
Q	y	2280	GAGCAGAGCTGCTGAGAAACGCCAGAAAGATGCCAGAGAGTGGCTTGTCAGATGACAAGA	2339
D	b	1599	GAGCAGAGCTGCTGAGAAACGCCAGAAAGATGCCAGAAAGATGGCTTGTCAGATGACAAGA	1658
Q	y	2340	AACCAATTCAGATGTTTTCACCTCCTAAAGGTTTCATCTCTCAAGATAGAAGCGCAGCAGAT	2399
D	b	1659	AACCAATTCAGATGTTTTCACCTCCTAAAGGTTTCATCTCTCAAGATAGAAGCGCAGCAGAT	1718
Q	y	2400	TTTTGAATAGTCTGTGCAGAAAAGCAGTAGTGTTGTCMAATCGACCCTCAAGCAGCAGATAG	2459
D	b	1719	TTTTGAATAGTCTGTGCAGAAAAGCAGTAGTGTTGTCMAATCGACCCTCAAGCAGCAGATAG	1778
Q	y	2460	TCTCCAAGATTGACAGCAGACTGGAGCAGTATACCAAGTGCAATTCAGGGAACAAAAGCG	2519
D	b	1779	TCTCCAAGATTGACAGCAGACTGGAGCAGTATACCAAGTGCAATTCAGGGAACAAAAGCG	1838
Q	y	2520	CAAAACCTACAAAGCCGCGAGCCTCGGATCTTCCTGTTCTGCTGAAAGTGTCGCAACA	2579
D	b	1839	CAAAACCTACAAAGCCGCGAGCCTCGGATCTTCCTGTTCTGCTGAAAGTGTCGCAACA	1898
Q	y	2580	TCAAGAGTATGTCGGAGAAAGGGAATGTTGTTTTCAATCCCCCCTCAGCAGGCAACACCAA	2639
D	b	1899	TCAAGAGTATGTCGGAGAAAGGGAATGTTGTTTTCAATCCCCCCTCAGCAGGCAACACCAA	1958
Q	y	2640	ATAAGGAAACTGCTGGCTTGAAGGTAGGGGTTTCTAGCCGCAATCAATGAATGGCTAACTA	2699
D	b	1959	ATAAGGAAACTGCTGGCTTGAAGGTAGGGGTTTCTAGCCGCAATCAATGAATGGCTAACTA	2018
Q	y	2700	AAACCCAGATGGAACAAGTCACCTGCTCCCAACCTTCTGACTTTGAGACCAGGAGCG	2759
D	b	2019	AAACCCAGATGGAACAAGTCACCTGCTCCCAACCTTCTGACTTTGAGACCAGGAGAGCG	2078
Q	y	2760	TATCCAGCAGCGGAACTCTGGGAAAAGCAATCTGTGGATAAGGTCACTTCCCCCACTA	2819
D	b	2079	TATCCAGCAGCGGAACTCTGGGAAAAGCAATCTGTGGATAAGGTCACTTCCCCCACTA	2138
Q	y	2820	AGGTTTGAACAGTTTCAGAAAAGAACCAAGCTCAAGACGCAAGCAGCAGCTCAGTTGTAG	2879
D	b	2139	AGGTTTGAACAGTTTCAGAAAAGAACCAAGCTCAAGACGCAAGCAGCAGCTCAGTTGTAG	2198

Qy	2880	AGGGCTAAATTCGCTCTGTGTTTGGTAATATAGTTGAAATTTACTTAAATTTGGGTTCAATATCTTTT	29339
Db	2199	AGGGCTAAATTCGCTCTGTGTTTGGTAATATAGTTGAAATTTACTTAAATTTGGGTTCAATATCTTTT	2258
Qy	2940	TATTTTTCAAATCCAGTAAACCCCATGATATATTACATATATTTAATTAATACACAGTCT	2999
Db	2259	TATTTTTCAAATCCAGTAAACCCCATGATATATTACATATATTTAATTAATACACAGTCT	2318
Qy	3000	AGAGATGTTTCATGGTAAAGTAAGTACTGCTTTTGACAGAGAGCCCTGTTTCTAAAGAAACCCCAT	3059
Db	2319	AGAGATGTTTCATGGTAAAGTAAGTACTGCTTTTGACAGAGAGCCCTGTTTCTAAAGAAACCCCAT	2378
Qy	3060	GCTGTGAAATACAGACTTTTCTACTGATCATATAAATCTCTGTATCTGAGCAGTGATACCA	3119
Db	2379	GCTGTGAAATACAGACTTTTCTACTGATCATATAAATCTCTGTATCTGAGCAGTGATACCA	2438
Qy	3120	ACCACATCTGAAGTCAACAGAGAGATCCAAGTTTAAAAATTTGCCTCGGGAATGTGTGCAGTA	3179
Db	2439	ACCACATCTGAAGTCAACAGAGAGATCCAAGTTTAAAAATTTGCCTCGGGAATGTGTGCAGTA	2498
Qy	3180	TCTAGAAAAATGAAACCGTAGTTTTTGTTTTTTTTAAATACAGAAGTCATGTTCTTTCTGCA	3239
Db	2499	TCTAGAAAAATGAAACCGTAGTTTTTGTTTTTTTTAAATACAGAAGTCATGTTCTTTCTGCA	2558
Qy	3240	CTTTTAAATAAAGCATGGAAGAAATATCTTAGTAGGCAATTTGTAACACTTTTTTGAAGAGT	3299
Db	2559	CTTTTAAATAAAGCATGGAAGAAATATCTTAGTAGGCAATTTGTAACACTTTTTTGAAGAGT	2618
Qy	3300	AACCCATTTACAGATTTGAAATACCTCGGATAAATGGTTGCTTTTAAAAAAAAGAAAAGATG	3359
Db	2619	AACCCATTTACAGATTTGAAATACCTCGGATAAATGGTTGCTTTTAAAAAAAAGAAAAGATG	2678

RESIN, T 8

```

US-10-062-674-1816
; Sequence 1816, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 03/625,102
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PERL Program
; SEQ ID NO 1816
; LENGTH: 4261
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040005559A1 256147.57
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (4261)
; OTHER INFORMATION: a, t, c, g, or other
US-10-062-674-1816

```

Query Match	47.8%	Score 1771;	DB 15;	Length 4251;
Best Local Similarity	76.7%	Pred. No. 0;		
Matches 2578:	Conservative	0: Mismatches	16:	Indels 768:
				Gaps 3:

Qy	1	GCCTAGGGAATGA	CAGGCATCTCCACAGGCAGGCTGCATCCACCTT	-GGCTGGGGTGTGG	59
Db	84	GCCTAGGGAATGA	CAGGCATCTCCACAGGCAGGCTGCATCCACCTTGGGCT	GGGGTGTGG	143
Qy	60	TCATTGGCTGCCTATTAGAAAAACGACAGCAATGCATACACCGCCTCCCGACTGTAA			119
Db	144	TCATTGGCTGCCTATTAGAAAAACGACAGCAATGCATACACCGCCTCCCGACTGTAA			203

Db	849	-----	848
Qy	1597	AGGCAGAAACGTAAACAAGCAGCTAGAGAGAAAAAACAATGCCATGCAAGAGACAAAGAT	1656
Db	849	-----	848
Qy	1657	AAAAGGGAAAGGTAGAACAGAAATAGAAGGGAATGGTAATGTAATGAAAGAAAGCACA	1716
Db	849	-----	848
Qy	1717	AGAAGATAAATTACAGACAGCTGCTCTAAAGAAACAGGGAGACAGAGAAGGAACTAAAGT	1776
Db	849	-----	848
Qy	1777	GCAAGCTAAACAGAAAAAGCTCCAAGAGACAAGCCCTACTCTTCAAAAAAGAGAGATCAA	1836
Db	849	-----	855
Qy	1837	AGATGAAAAGATTAAAAAGGACAAAGAACCCAAAGAAAGAGTTAAGAGCTTCATGGATCG	1896
Db	856	AGATGAAAAGATTAAAAAGGACAAAGAACCCAAAGAAAGAGTTAAGAGCTTCATGGATCG	915
Qy	1897	AAAGAAAGGATTACAGAAAGTTAAGTCGCAGAATGGAGAAATTCATGCCCAAACTTAA	1956
Db	916	AAAGAAAGGATTACAGAAAGTTAAGTCGCAGAATGGAGAAATTCATGCCCAAACTTAA	975
Qy	1957	ACATACTGAGATACTTTACGCCCTTGAGGAGGGCCAGCTGGACACCAAGGAGGC	2016
Db	976	ACATACTGAGATACTTTACGCCCTTGAGGAGGGCCAGCTGGACACCAAGGAGGC	1035
Qy	2017	TGAGGGCGCCCGCCAGGTGGAAGCCGCAAAAGGCTGGAGAGAGTTCTGCTGTCGCGG	2076
Db	1036	TGAGGGCGCCCGCCAGGTGGAAGCCGCAAAAGGCTGGAGAGAGTTCTGCTGTCGCGG	1095
Qy	2077	GGAGACCGAGCGGAAGAGTTTCAGAGAGCTCAAAACAGNAGCAGCAGGAGGGCTTTGGA	2136
Db	1096	GGAGACCGAGCGGAAGAGTTTCAGAGAGCTCAAAACAGNAGCAGCAGGAGGGCTTTGGA	1155
Qy	2137	GCTGAGGAACTCAAGAAAAGAGGAGAGAGAAAGAGTCTCTGGAGGAGAAAGACA	2196
Db	1156	GCTGAGGAACTCAAGAAAAGAGGAGAGAGAAAGAGTCTCTGGAGGAGAAAGACA	1215
Qy	2197	GAGGAGGACGAGGAGGAGCCGATCAAAACTCAGAGAGGAGGAGAGAGAGGAGGCT	2256
Db	1216	GAGGAGGACGAGGAGGAGCCGATCAAAACTCAGAGAGGAGGAGAGAGAGGAGGCT	1275
Qy	2257	AAAGGAAGAGATTGAAAGGCGAAGAGCAGAGCTGCTGAGAAACCCAGAAAGATCCAGA	2316
Db	1276	AAAGGAAGAGATTGAAAGGCGAAGAGCAGAGCTGCTGAGAAACCCAGAAAGATCCAGA	1335
Qy	2317	AGATGGCTTCAGATGACAGAAACCATCAAGTGTTTCACTCCTCAAGGTTTCACTCT	2376
Db	1336	AGATGGCTTCAGATGACAGAAACCATCAAGTGTTTCACTCCTCAAGGTTTCACTCT	1395
Qy	2377	CAAGATAGAGAGCGACAGAAATTTTGAATAGTCTGTGCAAGAAAAGCAGTGGTCAA	2436
Db	1396	CAAGATAGAGAGCGACAGAAATTTTGAATAGTCTGTGCAAGAAAAGCAGTGGTCAA	1455
Qy	2437	ATCGACCCATCAGACGCAATAGTCTCAAGATTGACAGCAGCTGGAGCAGTATACCAG	2496
Db	1456	ATCGACCCATCAGACGCAATAGTCTCAAGATTGACAGCAGCTGGAGCAGTATACCAG	1515
Qy	2497	TGCAATTGAGGGAAACAAAAAGCGCAAAACCTACAAAGCGCGCCTCGATCTTCCTGT	2556
Db	1516	TGCAATTGAGGGAAACAAAAAGCGCAAAACCTACAAAGCGCGCCTCGATCTTCCTGT	1575
Qy	2557	TCCTGCTGAAGGTGATCGCAACATCAAGAGTATGTGGGAGAAAGGAATGTGTTTTCA	2616
Db	1576	TCCTGCTGAAGGTGATCGCAACATCAAGAGTATGTGGGAGAAAGGAATGTGTTTTCA	1635
Qy	2617	CCCCACTGCAGCGGACACCAAAATAAGAAACTGCTGGCTTGAAGGTAGGGTTTCTAG	2676
Db	1636	CCCCACTGCAGCGGACACCAAAATAAGAAACTGCTGGCTTGAAGGTAGGGTTTCTAG	1695

Qy	2677	CCGATCAATGATGGCTAACTAATAAACCCAGAGTCACTGCTCCAAACC	2733
Db	1696	CCGATCAATGAATGGCTAACTAATAAACCCAGAGTCACTGCTCCAAACC	1755
Qy	2737	TTCTGACTTGAGACAGGAGACGTATCCAGCAAGCGAAACCTCTGGGAAAGCAATCTGT	2796
Db	1756	TTCTGACTTGAGACAGGAGACGTATCCAGCAAGCGAAACCTCTGGGAAAGCAATCTGT	1815
Qy	2797	GGATAAGGCTCACTTCCCCACATAAGGTTTGAGACAGTTCCAGAAAGAACCCAGGCTCAAG	2856
Db	1816	GGATAAGGCTCACTTCCCCACATAAGGTTTGAGACAGTTCCAGAAAGAACCCAGGCTCAAG	1875
Qy	2857	ACGCAGGACGAGCTCAGTTGTGAGGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTT	2916
Db	1876	ACGCAGGACGAGCTCAGTTGTGAGGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTT	1935
Qy	2917	ACTAAATTTGGGTTTCATTTATCTTTTATTTTTTCAATATCCCCAGTAAACCCATGTATATTATC	2976
Db	1936	ACTAAATTTGGGTTTCATTTATCTTTTATTTTTTCAATATCCCCAGTAAACCCATGTATATTATC	1995
Qy	2977	ACTATATTTAATAANTCACAGCTCAGAGATGTTTCATGCTAAGAGTACTGCCCTTTGCACAGS	3036
Db	1996	ACTATATTTAATAANTCACAGCTCAGAGATGTTTCATGCTAAGAGTACTGCCCTTTGCACAGS	2055
Qy	3037	AGCCTGTTTCTAAAGAAACCCATGCTCTGTAAATAGAGACTTTTCTACTGATCATCATAAAC	3096
Db	2056	ATCCTGTTTCTAAAGAAACCCATGCTCTGTAAATAGAGACTTTTCTACTGATCATCATAAAC	2115
Qy	3097	TCTGTATCTGACGAGTATACCAACCAACATCTGAAGTCAACAGAGAATGCCAAGTTTAAAA	3156
Db	2116	TCTGTATCTGACGAGTATACCAACCAACATCTGAAGTCAACAGAGAATGCCAAGTTTAAAA	2175
Qy	3157	TTGCTCGGGAATGHTGTCAGTATCTAGAAAAAATGAACCGTAGTTTTTGTGTTTTTAAAT	3216
Db	2176	TTG-CTGCGGGAATGHTGTCAGTATCTAGAAAAAATGAACCGTAGTTTTTGTGTTTTTAAAT	2234
Qy	3217	ACAGAAGTCATGTTGTTTCTGCACTTTATATAAAGCATGGAAGAAATTTATCTTAGTAGG	3276
Db	2235	ACAGAAGTCATGTTGTTTCTGCACTTTATATAAAGCATGGAAGAAATTTATCTTAGTAGG	2294
Qy	3277	CAATTTGAACACTTTTTTGAAGTAACCCATTCAGATTTGAAATCTCGGATAATCGTTG	3336
Db	2295	CAATTTGAACACTTTTTTGAAGTAACCCATTCAGATTTGAAATCTCGGATAATCGTTG	2354
Qy	3337	CTTTAAAAAAGAAAAAGATCGGCGACAGTTTATTCCTTT	3378
Db	2355	CTTTAAAAAAGAAAAAGATGTAAGTTGTAAGGTATTAATTT	2396

RESULT 11

```

US-09-765-231A-32
; Sequence 32, Application US/09765231A
; Patent No. US20020119452A1
; GENERAL INFORMATION:
; APPLICANT: Searle/Monsanto
; APPLICANT: Phlippard, Deborah
; APPLICANT: Vasanthakumari, Geetha
; APPLICANT: Dotson, Stanton
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Osteoarthritis tissue-
; TITLE OF INVENTION: vectors, and cells
; FILE REFERENCE: SO-3221 PR
; CURRENT APPLICATION NUMBER: US/09/765,231A
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 82
; SEQ ID NO 32
; LENGTH: 3342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-765-231A-32

```

Query Match 39.3%; Score 1455.8; DB 9; Length 3342;

Best Local Similarity 99.4%; Pred. No. 0; Matches 1482; Conservative 0; Mismatches 7; Indels 2; Gaps 2;									
QY	1871	GAAGAAAGTTAAGAGCTTCATGGATCGAAGAGGATTTACAGAAAGTTAAAGTCGAGAAAT	1930						
DB	1	GAAGAAGTTAAGAGCTTCATGGATCGAAGAGGATTTACAGAAAGTTAAAGTCGAGAAAT	60						
QY	1931	CGAGAATTCATGACCCACAACTTAACATACACTAGAAATATTTTCAGCGCCCTGGAGGG	1990						
DB	61	CGAGAATTCATGACCCACAACTTAACATACACTAGAAATATTTTCAGCGCCCTGGAGGG	120						
QY	1991	AGGGCCAGCGTGCACACCAAGGAGCTGAGGGCCGCCCGAGGTGGAAGCGGCAAAAGG	2050						
DB	121	AGGGCCAGCGTGCACACCAAGGAGCTGAGGGCCGCCCGAGGTGGAAGCGGCAAAAGG	180						
QY	2051	CTGGAGAGCTTCGTCTGCTCGCGGGAGACCGAGAGCGAAGAGTTCGAGAAAGTCTCAA	2110						
DB	181	CTGGAGAGCTTCGTCTGCTCGCGGGAGACCGAGAGCGAAGAGTTCGAGAAAGTCTCAA	240						
QY	2111	CAGAAGCAGCAGAGCGGCTTTGGAGCTGGAGAACTCAGAGAAACAGAGGAGGAGAGA	2170						
DB	241	CAGAAGCAGCAGAGCGGCTTTGGAGCTGGAGAACTCAGAGAAACAGAGGAGGAGAGA	300						
QY	2171	AGGAAGCTCTCGAGAGGAGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2230						
DB	301	AGGAAGCTCTCGAGAGGAGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	360						
QY	2231	AGAGAGGAGAGAGAGAGAGAGGCTTAAGAGAGAGATTTGAAGCGAGAGCAGAGAGCT	2290						
DB	361	AGAGAGGAGAGAGAGAGAGAGGCTTAAGAGAGAGATTTGAAGCGAGAGCAGAGAGCT	420						
QY	2291	GCTGAGAAACCCAGAGAGATGCCAGAGATGGCTTGTGAGATGACAGAAACCAATTCAG	2350						
DB	421	GCTGAGAAACCCAGAGAGATGCCAGAGATGGCTTGTGAGATGACAGAAACCAATTCAG	480						
QY	2351	TGTTTCATCTCTAAGAGTTCATCTCTCAAGATAGAGAGAGAGAGAGATTTTGAATAG	2410						
DB	481	TGTTTCATCTCTAAGAGTTCATCTCTCAAGATAGAGAGAGAGAGATTTTGAATAG	540						
QY	2411	TCGTGCAGAAAGCAGTGTGTCAATCGACCATCAAGCAGCAATAGTCTCAAGATT	2470						
DB	541	TCGTGCAGAAAGCAGTGTGTCAATCGACCATCAAGCAGCAATAGTCTCAAGATT	600						
QY	2471	GACAGCAGACTGGAGCAGTATACCAAGTGCATTTAGGGAAACAAAGCGCAAAACCTACA	2530						
DB	601	GACAGCAGACTGGAGCAGTATACCAAGTGCATTTAGGGAAACAAAGCGCAAAACCTACA	660						
QY	2531	AGCGCGCAGCTCGATCTCTCTCTCTGCTGAGGTGTACGCAACATCAAGAGTATG	2590						
DB	661	AGCGCGCAGCTCGATCTCTCTCTCTGCTGAGGTGTACGCAACATCAAGAGTATG	720						
QY	2591	TGGGAGAAAGGGAATGTGTTTTCATCCCCCACTGCAGCAGGCACACCAAAATAAGGAACT	2650						
DB	721	TGGGAGAAAGGGAATGTGTTTTCATCCCCCACTGCAGCAGGCACACCAAAATAAGGAACT	780						
QY	2651	G-CTGGCTTGAAGTAGGGTTCTAGCCGATCAATGAATGGCTTAATAAACCACCA	2709						
DB	781	GCTGGCTTGAAGTAGGGTTCTAGCCGATCAATGAATGGCTTAATAAACCACCA	840						
QY	2710	TGGAAACAGTCACTGCTCCCAACCTTCTGACTTGAGACAGGAGAGCTATCCAGCAA	2769						
DB	841	TGGAAACAGTCACTGCTCCCAACCTTCTGACTTGAGACAGGAGAGCTATCCAGCAA	900						
QY	2770	CGGAAACCTCTGGAAAGCAATCTGTGGATAAGGTCACTTCGCCCACTAAGTTTGA	2828						
DB	901	CGGAAACCTCTGGAAAGCAATCTGTGGATAAGGTCACTTCGCCCACTAAGTTTGA	960						
QY	2829	ACAGTTCAGAGAAACCCAGAGCTCAAGCAGCAGGAGCTCAGTTGTAGAGGCTAAT	2888						
DB	961	ACAGTTCAGAGAAACCCAGAGCTCAAGCAGCAGGAGCTCAGTTGTAGAGGCTAAT	1020						
QY	2889	TCGCTCTGTTTGTGATTTATGTTGATTTACTAAATTTGGGTTTCATTTATTTTCA	2948						

RESULT 12
US-09-925-301-100
; Sequence 100, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 2511
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12)-
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (28)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (44)-
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (2456)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (2488)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (2511)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-100
Query Match 15.8%; Score 594.2; DB 9; Length 2511;

[illegible]

```

; SEQ ID NO 35
;
; LENGTH: 2209
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; FEATURE:
;
; NAME/KEY: misc feature
;
; OTHER INFORMATION: Incyte ID No. US20020156263A1 022404.23
US-09-974-298-35

```

Query Match	14.8%	Score 546.8;	DB 9;	Length 2209;
Best Local Similarity	78.3%	Pred. No. 6.2e-129;		
Matches 808;	Conservative	0;	Mismatches 7;	Indels 217;
				Gaps 4;

Qy		2328	CAGN	GACAGA	AACCA	TTCNA	GTGTTT	CACTC	CTAA	AGGTT	CA	TCTCT	CA	AGATGA	GAAG	2387																			
Dd		1	CAGAT	GACAGA	AACCA	TTCNA	GTGTTT	CACTC	CTAA	AGGTT	CA	TCTCT	CA	AGATGA	GAAG	60																			
Qy		2388	AGCG	AGCAGA	ATTTTT	TGA	ATAA	GTC	TG	GCAG	AAAA	CAG	TGG	TGTC	CAAA	TGCA	CCCATC	2447																	
Dd		61	AGCG	AGCAGA	ATTTTT	TGA	ATAA	GTC	TG	GCAG	AAAA	CAG	TGG	TGTC	CAAA	TGCA	CCCATC	120																	
Qy		2448	AAGC	AGCAA	TAGT	CTCC	ACA	GATT	CA	CAG	CAG	CAG	CTG	GAG	CAG	TAC	CAG	TG	CAAT	TCAGG	2507														
Dd		121	AAGC	AGCAA	TAGT	CTCC	ACA	GATT	CA	CAG	CAG	CAG	CTG	GAG	CAG	TAC	CAG	TG	CAAT	TCAGG	180														
Qy		2508	GA	CAAAA	AAG	CGCA	AAAA	CC	TACA	AA	AG	CG	CG	CA	GCT	CG	GA	TCT	GT	TCT	GT	CT	CT	GAAG	2567										
Dd		181	GA	CAAAA	AAG	CGCA	AAAA	CC	TACA	AA	AG	CG	CG	CA	GCT	CG	GA	TCT	GT	TCT	GT	CT	CT	GAAG	240										
Qy		2568	GT	GT	PAC	CA	CA	CAT	CA	AG	GAT	AT	GT	GG	GA	AA	AG	GA	AT	GT	TTTT	CA	T	CCCC	CA	CT	CG	AG	2627						
Dd		241	GT	GT	PAC	CA	CA	CAT	CA	AG	GAT	AT	GT	GG	GA	AA	AG	GA	AT	GT	TTTT	CA	T	CCCC	CA	CT	CG	AG	300						
Qy		2628	CAG	CA	CA	CA	AA	TAA	GG	AA	ACT	CT	CG	CT	GT	TAA	AG	GT	AG	GG	TTTT	CT	AG	CG	CG	CA	TCA	AT	GT	2687					
Dd		301	CAG	CA	CA	CA	AA	TAA	GG	AA	ACT	CT	CG	CT	GT	TAA	AG	GT	AG	GG	TTTT	CT	AG	CG	CG	CA	TCA	AT	GT	360					
Qy		2688	AA	TGG	CT	TA	A	CT	TA	AA	A	CC	CC	CA	GT	AG	CA	CT	CA	CT	GT	CT	CC	CA	AA	CC	TT	CT	GAC	T	TTGA	2747			
Dd		361	AA	TGG	CT	TA	A	CT	TA	AA	A	CC	CC	CA	GT	AG	CA	CT	CA	CT	GT	CT	CC	CA	AA	CC	TT	CT	GAC	T	TTGA	420			
Qy		2748	GACC	GAG	AC	GT	AT	CC	AG	CA	AG	GA	CC	CT	CT	GG	AA	AA	GC	AT	CT	GT	GG	AT	AA	GC	AT	CT	GT	GG	AT	AA	GC	AT	2807
Dd		421	GACC	GAG	AC	GT	AT	CC	AG	CA	AG	GA	CC	CT	CT	GG	AA	AA	GC	AT	CT	GT	GG	AT	AA	GC	AT	CT	GT	GG	AT	AA	GC	AT	480

2808	Qy	CTTCCCCCACTAAGGTTTGAGACAGATTCCAGAAAGAAACCCCAAGCTCAAGACGCAGGACGA	288
481	Db	CTTCCCCCACTA	492
2568	Qy	GCTCAGTTGTAGAGGGCTAATTCGCTCTGTTTGTATTATATGTTGATTACTAAATTTGGG	2927
493	Db	-----	492
2528	Qy	TTCATTATCTTTTATTTTCAATATCCCAAGTAAACCCCATGTTATTTATCTACTATATTAA	2987
493	Db	-----	492
2988	Qy	TAATCAGCTCTAGAGATGTTTCATGGTAAAGGTACTGCCTTTGACAGGAGCGCTGTTTCT	3047
493	Db	-----AGGAGCCCTGTTTCT	506
3048	Qy	AAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCATAACTCTGTATCTGA	3107
507	Db	AAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCATAACTCTGTATCTGA	566
3108	Qy	GCAGTGATACCAACCAACCATCTGGAAGTCAACAGAGAGATCCCAAGTTTAAAAATTGCCTCGCGA	3167
567	Db	GCAGTGATACCAACCAACCATCTGGAAGTCAACAGAGATCCCAAGTTTAAAAATTGCCTCGCGA	626
3168	Qy	ATGTGTGCAGTATCTAGAAAATGAAACCGTAGTTTTTTTGTTTTTTTTAAATACAGAAGTCAT	3227
627	Db	ATGTGTGCAGTATCTAGAAAATGAAACCGTAGTTTTTTTGTTTTTTTTAAATACAGAAGTCAT	686

```

QY 3328 GTTCTTTTCGACATTATATAATTAAGCATGGAAGAAATATCTTTAGTAGGCATTGTAACA 3328
Db 687 GTTGTTCCTCGACATTATATAATTAAGCATGGAAGAAATATCTTTAGTAGGC-ATTGTAACA 745
QY 3328 CTTTTTGAAGTAACCCCATTTTCAGATTTCGATAATTCGATAATGGTTGCTTTTAAAAAAA 3347
Db 746 C-TTTTGAAGTAACCCA-TTCAGATTTCGATAATGCCATAATGGTTGCTTTTAAAAAAA 803
QY 3348 AAGAAAAGATG 3359
Db 804 AAAAAAGAAATG 815

RESULT 15
US-10-085-783A-16172
; Sequence 16172, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16172
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-16172

```

Query Match	13.5%;	Score 499;	DB 12;	Length 541;
Best Local Similarity	96.2%;	Pred. No. 4.1e-117;		
Matches 511; Conservative	0;	Mismatches 20;	Indels	Gaps 0;
2848	AAGCTCAAGACGACGAGACGAGCTCAGTGTAGAGGGCTAAATCGCTCTGTTTGTATTATTA	2907		
Db	1 AAGCTCAGACGACGAGACGAGCTCAGTGTAGAGGGCTAAATCGCTCTGTTTGTATTATTA	60		
2908	TGTTGATTTTACTAAATTTGGGTTTCATTATCTTTTATTTTTCAAATATCCAGTAAACCCATG	2967		
Db	61 TGTTGATTTTACTAAATTTGGGTTTCATTATCTTTTATTTTTCAAATATCCAGTAAACCCATG	120		
2968	TATATTATCATTATATTTAAATATCAACAGTCTAGAGATGTTTCATGGTAAAAAGTACTGCGCT	3027		
Db	121 TATATTATCATTATATTTAAATATCAACAGTCTAGAGATGTTTCATGGTAAAAAGTACTGCGCT	180		
3028	TTGCACAGAGCGCTGTTTCTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGAT	3087		
Db	181 TTGCACAGAGCGCTGTTTCTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGAT	240		
3088	CATCATTACTCTGTATCTGAGCAGTGTATACCAACCATCTGGAAGTCAACAGAGAATCCA	3147		
Db	241 CATCATTACTCTGTATCTGAGCAGTGTATACCAACCATCTGGAAGTCAACAGAGAATCCA	300		
3148	AGTTTAAATTTGCCTCGCGAATGTGTGCAGTATCTAGAAAAATCAACCGTAGTTTTTGTT	3207		
Db	301 AGTTTAAATTTGCCTCGCGAATGTGTGCAGTATCTAGAAAAATCAACCGTAGTTTTTGTT	360		
3208	TTTTTAAATACAGAGTCAATGTTTCTGCACTTTTATATAAAGCATGGAAGAAATTAAT	3267		
Db	361 TTTTTAAATACAGAGTCAATGTTTCTGCACTTTTATATAAAGCATGGAAGAAATTAAT	420		
3268	CTTAGTAGGCAATTGTAACACTTTTTTGAAGTAACCCATTTTCAGATTTGAAATTAATCTGCGA	3327		
Db	421 CTTAGTAGGCAATTGTAACACTTTTTTGAAGTAACCCATTTTCAGATTTGAAATTAATCTGCGA	480		

